How to improve the design of experimental studies in computing education: Evidence from the international assessments

Mustafa Yildiz^{1,2} · Ran Shi³ · Mehmet Kara^{4,5}

Received: 19 August 2021 / Accepted: 2 November 2021 / Published online: 13 January 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Cluster randomized trials are frequently used in educational research for methodological reasons. This study aims to improve the efficiency of cluster randomized trials on computer/information literacy and computational thinking. The study employs a two-level hierarchical linear model to estimate (i) intraclass correlation coefficients, (ii) the amount of explained variances given selected predictors, and (iii) minimum detectable effect sizes given the set of plausible scenarios. Two data cycles from the International Computer and Information Study were used. The covariates at the student level are gender, interest in ICT, parents' highest education level, ICT self-efficacy, and experience with computers. The covariates at school/teacher level are teacher's ICT use, ratio of school size to the number of computers for student use, availability of ICT resources at school, approximate teacher age, and ICT self-efficacy. Findings showed that the most precise effect could be measured when student and teacher/school covariates are both adopted. Lastly, it was revealed that increasing the number of schools is effective to get the most precise effect.

Keywords Computer and information literacy · Computational thinking · Cluster randomized trials · Power analysis · ICILS

1 Introduction

With the advancement in technology during the past few decades, the availability of Information and Communication Technologies (ICT) to human use has flourished. This change influenced educational practices in a way that ICT use has become part of teaching and learning practices, and so, often schooling systems have made efforts to adjust to the current situation by having students and teachers develop relevant skills.

✉ Mehmet Kara
m.kara@live.com

Extended author information available on the last page of the article

Given the current conditions started by COVID-19, many educational institutions worldwide have converted their mode of teaching from face-to-face to online learning. This necessitated students and teachers to go online for educational needs. This change forced teachers and students to develop skills to get better ICT-related literacies. Therefore, the importance of the investigations with regard to ICT-related literacies in education is even more critical today.

The incorporation of ICT into education took place for a long time, while the conceptual understanding of ICT-use has kept evolving as ICT itself has been changing rapidly along the way. Given the current conditions started by COVID-19, many institutions worldwide have converted their mode of teaching from face-to-face to online and technology integration has gained more importance. Some of the research areas related to ICT in terms of embedding technology into educational practices are media literacy (Livingstone & Van der Graaf, 2008), digital literacy Bawden, 2008), computer literacy (Hoffman & Blake, 2003), information literacy (American Library Association, 2000), twenty-first-century skills (Dede, 2010), digital competence (Ferrari, 2012), and computational thinking (Wing, 2006). The skillset within each conceptualization, with some overlaps, differs such that being able to use computers/devices effectively (computer literacy), being able to locate, access, and process information (information literacy; American Library Association, 2000), to have the skills such as information management, critical thinking, creativity, collaboration, problem-solving (twenty-first-century skills; van Laar et al., 2017), to use computers to create solutions to real-life issues by decomposition, abstraction, algorithm design, debugging, iteration and generalization (computational thinking; Shute et al., 2017). The motivating idea behind these research endeavors is that we live in a society with an economic and social structure that necessitates individuals to develop relevant skills to effectively and successfully participate. As a result, enhancing education in terms of technology-related skills becomes essential in today's technology-driven world.

There have been increasing efforts to measure and improve students' ICT-related skills over the years in this context. One of the most significant efforts is an international assessment started in 2013 by the International Association for Evaluation of Educational Achievement (IEA), which is called the International Computer and Information Literacy Study (ICILS).

ICILS studies have provided evidence and implications for the improvement of efforts in learners' ICT-related competencies. In this sense, these studies also have a potential to improve the design of experimental studies in ICT-related fields. From a methodological perspective, Cluster Randomized Trials (CRTs), a type of experimental research design, are considered the golden standard for educational research (Torgerson & Torgerson, 2001). In this study, we aim to improve the design of CRTs that are to be conducted on CIL and/or CT constructs to determine the design parameters such as intraclass correlation coefficients, amount of variance explained, and minimum detectable effect sizes. We hypothesize that countries would differ in terms of between-school and within-school variability, and thus, the CRTs to be designed within a specific country would depend on its contextual factors such as student demographics, teacher characteristics, and school characteristics (Fig. 1).

Countries participated in ICILS 2013 or 2018

Fig. 1 The list of countries that participated in an ICILS study at least once

2 Theoretical framework

2.1 Computer and information literacy

Computer and information literacy (CIL) is partly derived from the prior research that attempted to characterize ICT-related concepts in educational research such as digital literacy, ICT literacy, and information literacy, etc. According to Ng (2012), digital literacy was characterized as a construct referring to the use of online and offline digital tools for learning. Lau and Yuen (2014) described ICT literacy as a multidimensional construct that encompasses information literacy, internet literacy, and computer literacy. In the ICILS framework, CIL was proposed as a multidimensional construct, including both computer and information literacy, which was defined as “an individual’s ability to use computers to investigate, create, and communicate in order to participate effectively at home, at school, in the workplace, and in society” (Fraillon et al., 2019, p. 16). The major difference between ICT literacy and CIL is the emphasis on computers in the definition of CIL. Unlike ICT literacy, encompassing the broader context of all digital tools, the focus of CIL is personal computers, including traditional computers and other devices with external keyboard and mouse (Fraillon et al., 2019).

In this perspective, ICILS studies have been one of the most significant steps forward in assessing CIL skills globally. The ICILS studies had two cycles so far, with 21 and 14 participant countries in 2013 and 2018, respectively. The purpose of the ICILS study was to prepare a framework that would help to discover how young students (8th graders) may improve their CIL-related skills (Fraillon, Schulz, Friedman, Ainley, & Gebhardt, 2015). The ICILS studies attracted researchers’ interest on the topic, and European Educational Research Journal devoted a special issue to ICILS 2013 due to the fact that it is the first of its kind being an assessment of CIL

acquisition (Fraillon et al., 2014). In the ICILS studies, CIL was initially defined as a multidimensional construct with two dimensions having the following specified sub-dimensions:

1. “Collecting and managing information: Knowing about and understanding computer use, accessing and evaluating information, managing information.”
2. “Producing and exchanging information: Transforming information, creating information, sharing information, using information safely and securely” (Fraillon & Ainley, 2010).

Although there are two dimensions of CIL, studies have shown high correlations between the two constructs (Fraillon et al., 2014; Bos et al., 2014, as cited in Gerick et al., 2017); hence, CIL was used as a unidimensional construct in ICILS studies (Gerick et al., 2017).

2.1.1 Correlates of CIL

A significant number of studies have been conducted to investigate the factors influential on learners’ CIL and ICT-related traits. The relevant literature presented in the following sections showed that student, teacher, and school characteristics are linked to the concepts like CIL and ICT literacy.

Self-efficacy. In terms of learner characteristics, it was first shown that ICT self-efficacy is a predictor of CIL (Hatlevik et al., 2018; Rohatgi et al., 2016). Rohatgi et al. (2016) pointed out that the self-efficacy of learners in basic ICT skills is positively related to CIL while self-efficacy in advanced ICT skills is negatively related. Self-efficacy also functions as a mediator between learners’ ICT use and CIL (Hatlevik et al., 2018; Rohatgi et al., 2016). Lastly, Gebhardt et al. (2019) demonstrated that learners’ interest in and enjoyment of ICT positively linked with their CIL.

Gender. Prior research investigating gender differences in CIL had contradictory findings, as some studies found that female learners had higher CIL scores than males (Aydin, 2021; Gebhardt et al., 2019; Hatlevik et al., 2018; Punter et al., 2017). In contrast, Rohatgi et al. (2016) found out no gender differences in relation to ICT use, self-efficacy, and CIL. A study by Alkan and Meinck (2016) pointed out that the moderating effect of gender in the link between ICT use and CIL was weak.

Experience with ICT. Learner experience of ICT use was also demonstrated as an influential factor in CIL achievement. Studies reported a positive relationship between ICT use and CIL scores (Alkan & Meinck, 2016; Rohatgi et al., 2016) as well as the computer experience and the CIL achievement (Aydin, 2021; Hatlevik et al., 2018). Alkan and Meinck (2016) further showed that students frequently using ICT for social communication demonstrated better performance in CIL.

School ICT resources. The second consideration in the evaluation of learners’ CIL is the availability of ICT resources at schools. Gerick (2018) discovered that school characteristics in some countries have significant influences on learners’ CIL achievement. Similarly, Drossel, Eickelman and Gerick et al. (2017) showed that the availability of ICT resources at schools significantly influences learners’ usage.

Additionally, Drossel et al. (2020) argued that students of the resilient schools have CIL scores above the average despite their low socio-economic status. Hatlevik et al. (2018) indicated that learner use and experience of ICT in schools are the significant predictors of their CIL. They, however, noted that the quality of ICT usage in classrooms deserves more attention than the frequency of usage.

Teacher characteristics. Teacher characteristics are also linked to the use of ICT use frequency for teaching, which might eventually influence student CIL scores. Teachers' ICT self-efficacy (Drossel et al., 2017b; Siddiq et al., 2016), attitude and beliefs (Drossel et al., 2017b; Eickelmann & Vennemann, 2017), attendance in professional development (Drossel & Eickelmann, 2017), and availability of pedagogical (Gerick et al., 2017) and technical support (Drossel et al., 2017a) are denoted as the influential factors on teachers' ICT use for teaching. Additionally, Eickelmann and Vennemann (2017) revealed that relatively younger teachers, mostly classified as ICT enthusiasts, reported a higher frequency of ICT usage in classes. Previous research showed that teachers' age (Drossel et al., 2017a; Drossel et al., 2017b; Eickelmann & Vennemann, 2017; Siddiq et al., 2016), and gender (Drossel et al., 2017a; Drossel et al., 2017b) are influential on their ICT usage for teaching while Gebhardt et al. (2019) found out no significant influence of gender.

2.2 Computational thinking

Computational Thinking (CT) has been commonly adopted as a set of problem-solving skills, and it has taken much attention since Wing's (2006) conceptualization. CT corresponds to solving real-life problems based on the fundamental concepts of computer science (Wing, 2006, 2008). However, it is not about just programming (Voogt et al., 2015); it is more about conceptualization or abstraction to solve problems, which is a combination of mathematical, engineering, and scientific thinking skills (Wing, 2006). Therefore, it is viewed as a twenty-first-century skill that students could acquire with or without the use of computers (Shute et al., 2017). CT is currently considered as a fundamental skill for all age groups and required in all disciplines (Kalelioglu et al., 2016). It has been, for instance, included in the competency standards for students and teachers by the International Society for Technology in Education (ISTE, 2016).

CT was included in ICILS 2018 assessment, for the first time, as a problem-solving competency. Although the consensus among CT's definitions is limited to some extent (Shute et al., 2017; Voogt et al., 2015), it is defined in the ICILS framework as the "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). This definition likewise underscores conceptualizing, abstraction, designing systems (algorithmic solutions), and implementation via computers. Within ICILS 2018 framework, CT covered two strands and five aspects (Fraillon et al., 2019). These strands are (1) "Conceptualizing problems" and (2) "Operationalizing solutions" (Fraillon et al., 2019, p. 28). The first strand covers three aspects: "knowing about and understanding digital systems", "formulating and

analyzing problems”, and “collecting and representing relevant data” while the second strand covers “planning and evaluating solutions” and “developing algorithms, programs, and interfaces” (Fraillon et al., 2019, p. 28).

2.2.1 Correlates of CT

Learner demographics. Learner demographics related to CT skills were addressed in several studies (Tikva & Tambouris, 2021). In a recent study, Guggemos (2021) showed that learner characteristics such as CT self-concept, reasoning skills, gender, computer literacy and computer usage intensity, and learning opportunities at school are the predictors of CT skills. In a similar study, Durak and Saritepeci (2018) found out that thinking styles, ICT experience, internet usage intensity, and gender predict students’ CT skills. In the CT model proposed by Tikva and Tambouris (2021) learner demographics and the cognitive and non-cognitive factors were associated with CT knowledge and skills.

Interventions used to Improve Computational Thinking. A lot of interventions were designed and implemented to improve learners’ CT skills. Those interventions differ from one another in terms of their purpose and design. For instance, Hsu et al. (2018) identified project-based learning, problem-based learning, cooperative learning, and game-based learning as the most commonly used instructional methods, Tikva and Tambouris (2021) listed modeling/simulation-related and scaffolding strategies, Tang et al. (2020) indicated game-based learning and collaboration as the main pedagogical approaches. Other review studies specified reinforcement, reflection, information processing, scaffolding (Lye & Koh, 2014), storytelling (Ilic et al., 2018; Lye & Koh, 2014; Zhang & Nouri, 2019), and lectures (Ilic et al., 2018; Shute et al., 2017) as the commonly used methods. As for the tools/environments, programming tools (Ching et al., 2018; Hsu et al., 2018; Ilic et al., 2018; Lye & Koh, 2014; Tang et al., 2020; Tikva & Tambouris, 2021), robotics (Ching et al., 2018; Hsu et al., 2018; Ioannou & Makridou, 2018; Noh & Lee, 2020; Tikva & Tambouris, 2021), games or game development tools (Ching et al., 2018; Hsu et al., 2018; Kuo & Hsu, 2020; Tang et al., 2020), animation/virtual reality-related tools (Ching et al., 2018) and instructional modules (Shute et al., 2017) are stated as the frequently used interventions to improve CT skills.

One of the most common research designs in CT research is experimental research with an increasing usage frequency (Ilic et al., 2018). Scholars conducted empirical studies to investigate the effect of these interventions at diverse educational levels (Jun et al., 2017; Noh & Lee, 2020; Pala & Mihçı Türker, 2021; Sung et al., 2017). The majority of the reviewed experimental studies used quasi-experimental research design, including static groups without random assignment (Garneli & Chorianopoulos, 2018, 2019; Kuo & Hsu, 2020; Tsai et al., 2017; Yadav et al., 2014) while some of them used the one-group experimental design with pre and post-test (Jaipal-Jamani & Angeli, 2017; Noh & Lee, 2020). On the other hand, the studies in which experimental and control groups were formed with random assignment were rarely observed (Cetin, 2016; Jun et al., 2017; Pala & Mihçı Türker, 2021; Sung et al., 2017). Similarly,

Hickmott et al. (2018), in their review of CT studies in math, concluded that none of the studies conducted experiments with random assignment.

3 Significance and purpose of the study

Many studies have implemented interventions to observe what works better for computer literacy (e.g., Varank, 2006), information literacy (e.g., Chen & Chengalur-Smith, 2015; Gross & Latham, 2013; Walton & Hepworth, 2011), and computational thinking (Ching et al., 2018; Hsu et al., 2018; Lye & Koh, 2014; Shute et al., 2017; Tang et al., 2020; Tikva & Tambouris, 2021). However, our examination of the design of the experiments revealed that the use of cluster randomized trials as an experimental design in this area is scarce. One of the reasons that the CRTs were not being used so often could be due to the fact that there is no available prior research evidence to support the design of such complicated experimental studies as CRTs. In this respect, our study aims to contribute to the relevant literature by improving the knowledge base for efficient and rigorous design of intervention studies, in the form of CRTs, on both CIL and CT. In order to achieve our goal, we used the ICILS datasets provided by IEA from 2013 and 2018 data cycles. We used three sets of predictors: student characteristics, teacher/school characteristics, and student and teacher/school characteristics. The details about the nature of our predictor sets and models are explained in the paper's methods sections. The research questions that we intend to answer are:

1. What are the ICCs for CIL across the countries that participated in ICILS 2013 and ICILS 2018?
2. What are the ICCs for CT across the countries that participated in ICILS 2018?
3. How much variance could be explained on CIL given the set of student and teacher/school level characteristics across the countries that participated in ICILS 2013 and ICILS 2018?
4. How much variance could be explained on CT given the set of student and teacher/school level predictors across the countries that participated in ICILS 2018?
5. What would be the minimum detectable effect sizes given the sets of student and teacher/school level predictors, and selected constraints on CIL across the countries that participated in ICILS 2013 and ICILS 2018?
6. What would be the minimum detectable effect sizes given the sets of student and teacher/school level predictors, and selected constraints on CT across the countries that participated in ICILS 2018?

4 Methodology

4.1 Sample and variables

The data selected in this study were from the 2013 and 2018 cycles of the International Computer and Information Literacy Study (ICILS). The 2013 sample was

approximately 60,000 students from 21 participating countries, whereas the 2018 sample was around 46,000 students from 14 participating countries. Eligibility criteria for participation in the study required that the students be at least 13.5 years old and at the 8th year of schooling. The study's sampling design was described as 'complex' due to the use of multiple sampling strategies, including stratification of schools and randomization within schools for teachers and students (Drossel et al., 2017a).

The study's 2013 cycle only provided CIL assessment results, whereas the 2018 cycle of the study assessed both CIL and CT achievements. In this study, two sets of covariates for the two cycles describing student and teacher/school characteristics were used. Student characteristics included gender, self-efficacy for information and communications technology (ICT) use, student's interest in ICT, computer experience, and socio-economic status. The covariates used at the teacher/school level included the ratio of the school size to the number of computers in the school, ICT resources available in the school, the approximate age of the teacher, the ICT use of the teacher, and the self-efficacy of the teacher with regard to ICT use. The dependent variables (CT and CIL) were continuous latent variables estimated using a 1-parameter item response theory model (Rasch, 1993).

4.2 Analysis

A priori power analysis was conducted to estimate the adequate sample size in a study in order to generate accurate results by avoiding type II errors and utilizing resources economically. Computation of power analyses for cluster randomized trials necessitates five quantities; the intraclass correlation coefficients (ρ), the amount of variance explained (R^2), the balance of the design (P), the sample sizes for the clusters (J), and the number of units within the clusters (N). Therefore, this study provided country-specific evidence produced by power analysis aiming to improve the design of the CRTs concentrating on the CIL or CT. The analysis section of this paper outlined the following: the overview of the models used, the sets of covariates selected, the calculation of ICCs and its standard errors, the amount of variances explained by the covariates, the minimum detectible effect sizes, and how analytical details (plausible values, weights, software etc.) were handled.

4.2.1 Multilevel approach

The rationale behind employing multilevel models is the clustering of units in some forms of groups such as classrooms, clinics, and families (Raudenbush & Bryk, 2002). In this respect, employing CRTs in educational research would generate data sets that have multilevel structure. Since the sampling design of ICILS, as stated in the technical report (Fraillon et al., 2020), involved a systematic clustering at the school level, a two-level multilevel model was considered as students nested in schools. Therefore, the specification of the multilevel model given the sampling design could be characterized as a null model.

Null model:

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2) \quad (1)$$

$$\beta_{0j} = \gamma_{00} + u_{0j}, \quad u_{0j} \sim N(0, \sigma_\beta^2) \quad (2)$$

where Y_{ij} is the value of the dependent variable for the i^{th} student who attends to the j^{th} school. β_{0j} stands for the intercept of the j^{th} school, and γ_{00} is the average intercepts across the schools within a specific country. Lastly, ε_{ij} and u_{0j} stand for student and school level residuals, respectively.

The null model was considered to be a two-level model without any covariates. Initially, a three-level approach was considered to be a better fit due to the nature of the ICILS datasets in which students were nested in teachers, and teachers were nested in schools. However, the teacher data failed to incorporate into the analysis as in a three-level model because the teacher data were not linked to the student data due to the sampling procedures. Leaving the teacher data unused would not be wise as teachers are the central elements of any educational system. As a result, a central tendency measure (e.g., mean or median) of teacher covariates within a school was used at the second level as if it was a school level covariate. Please see Table 1 for more details about the models and the variables we have used.

4.2.2 Models used

Three models would be introduced for explaining additional variances as Model 1, Model 2, and Model 3 (full model). Model 1 included student-level covariates, Model 2 included teacher/school level covariates, and Model 3 included both student and teacher/school level covariates; please see Table 1 for more details about the models and the variables that were used. In these models, intercepts were freely estimated while the slopes were fixed across schools. Equations 3 and 4 illustrated the three models used:

Models with covariates at different/both level(s):

$$Y_{ij} = \beta_{0j} + \sum_{k=1}^K \beta_{kj} X_{kij} + \varepsilon_{ij}, \quad (3)$$

$$\beta_{0j} = \gamma_{00} + \sum_{k=1}^K \gamma_{0k} \bar{X}_{kj} + u_{0j}, \quad (4)$$

In Eqs. 3 and 4, i represents students, j represents schools, k represents covariates, Y_{ij} represents outcome variables (CIL or CT) for the i^{th} student who attends the j^{th} school, β_{0j} represents the intercept for the j^{th} school, γ_{00} represents the average intercepts across the schools within a country, β_{kj} is the slope for the k^{th} predictor for school j .

4.2.3 Intraclass correlations

The null model was used to see if there were substantial similarities among the students who attend the same school in terms of their CIL and CT scores. To do that,

Table 1 Variable descriptions by year and model

Variable type	2013	2018	2018	Variable	Type	Level	Model
Dependent variables	CIL	CIL	CT	Sex of the participant	Continuous		
	S_SEX	S_SEX	S_SEX		Categorical	1	1 and 3
	S_INTRST	–	–	Interest and enjoyment in using ICT	Continuous	1	1 and 3
	S_NISB	S_NISB	S_NISB	National index of students' socio-economic background	Continuous	1	1 and 3
Student level independent variables	S_BASEFF	S_GENEFF	S_GENEFF	ICT self-efficacy or self-efficiency for basic/general skills	Continuous	1	1 and 3
	–	S_EXCOMP	S_EXCOMP	Students' computer experience in years	Ordinal	1	1 and 3
	T_USETCH	–	–	Use of ICT for teaching at school	Continuous	2	2 and 3
	C_RATSTD	C_RATSTD	C_RATSTD	Ratio of school size and number of computers available for students	Continuous	2	2 and 3
Teacher/school level independent variables	C_ICTRES	C_ICTRES	C_ICTRES	Availability of ICT resources at school	Continuous	2	2 and 3
	T_AGE	T_AGE	T_AGE	Approximate age of teacher	Ordinal	2	2 and 3
	–	T_CLASACT	T_CLASACT	Use of ICT for classroom activities	Continuous	2	2 and 3
	T_EFF	T_ICTEFF	T_ICTEFF	Teachers ICT self-efficacy	Continuous	2	2 and 3

intraclass correlation coefficients (ICC) were estimated for each country. ICC represents the proportion of the between-school variance to the total variance, which was calculated as $\rho = \sigma_{\beta}^2 / (\sigma_{\beta}^2 + \sigma_{\epsilon}^2)$ where ρ is the ICC, σ_{β}^2 is the between-school variance and σ_{ϵ}^2 is the within-school variance. The denominator, $(\sigma_{\beta}^2 + \sigma_{\epsilon}^2)$, in this equation was the total variance. Further, the uncertainty associated with the ICCs were calculated using Donner and Koval's (1982) method as displayed in Kelcey and Phelps (2013) as shown below:

$$SE(\rho) = \sqrt{2(1 - \rho)^2 [1 + (N - 1)\rho]^2 / N(N - 1)J} \quad (5)$$

where ρ is the ICC, J is the number of schools, and N is the number of students within a school.

4.2.4 Additional variances explained

In multilevel models, the practical utility of a model is often evaluated in a similar way to the linear regression but at each level separately. The following equations showed how the proportion of variance explained by the inclusion of the predictors was calculated:

$$R_{\epsilon}^2 = (\sigma_{\epsilon}^2 - \sigma_{\epsilon}^2) / \sigma_{\epsilon}^2 \quad (6)$$

$$R_{\beta}^2 = (\sigma_{\beta}^2 - \sigma_{\beta}^2) / \sigma_{\beta}^2 \quad (7)$$

The terms σ_{ϵ}^2 and σ_{β}^2 represented the residual terms for the conditional models at level 1 and level 2, respectively.

The amount of additional variance explained was calculated as the ratio of the reduction of unexplained variance to the total unexplained variance in a given level. Hence, R_{ϵ}^2 and R_{β}^2 were the proportion of variance explained at level 1 and level 2 by inclusion of the covariates.

4.2.5 Minimum detectable effect sizes

Minimum detectable effect size (MDES; Bloom, 1995) is a design coefficient representing the smallest effect that an intervention can detect. Therefore, smaller values of MDES indicate smaller effects that can be detected by an intervention, which would suggest a better accuracy of the results. In this study, MDES was calculated for a balanced design ($P=0.50$) with a power value of 0.80 for a two-tailed test. Given these assumptions, the MDES were calculated using Hedges and Hedberg's (2007) method.

$$MDES = \frac{M_{(J-2-K)}}{\sqrt{P(1-P)}} \sqrt{\frac{\rho(1-R_{\beta}^2)}{J} + \frac{(1-\rho)(1-R_{\epsilon}^2)}{NJ}} \quad (8)$$

In this equation, ρ is the ICC, R^2 is the proportion of variance explained by the covariates, N is the number of students who attend to a school, J is the number of schools (clusters) in a country, P (0.50) stands for representing a balanced design, and M is a design multiplier ($M \approx t_{\alpha/2} + t_{\beta}$ for $J-2-K$ degrees of freedom where α and β are type I and type II error rates, respectively). MDES for the specific instances of N (20, 40, 60) and J (10, 20, 30, 40) were calculated in order to estimate the efficiency of CRTs.

4.2.6 Plausible values, software, sampling weights, and syntax

The two variables, CIL and CT, had five plausible value estimates, which were randomly drawn with regard to the students' ability from their posterior distribution. The models mentioned above (Null model, Model 1, Model 2 and Model 3) were first estimated for each plausible value, then the variance components ($\sigma_{\beta}^2, \sigma_{\epsilon}^2$) were calculated within each model. Then variance components were averaged across five plausible values. The MDES were calculated given the ICCs, R^2 s, and previously mentioned constraints (N, J, P , power). R (R Core Team, 2020) was used for modeling the multilevel structure using 'WeMix' package (Bailey et al., 2020), which allowed to use sampling weights at each level separately. Lastly, MDESs were calculated using 'PowerUpR' package (Bulus et al., 2019).

5 Results

Descriptive statistics such as sample sizes, number of schools, average teacher age, average CIL scores, percentage of males, and the average of the selected independent variables within each country for the 2013 cycle were displayed in Table 2. It shows the average teacher ages were similar across countries due to the lack of variation as the approximate age of teachers was coded during data collection. CZE and DNK had the highest CIL averages; on the contrary, THA and TUR had the lowest CIL averages. TUR and THA had the lowest CIL scores with the highest standard deviations (greatest variability on CIL). CZE was the highest performing country with an average CIL score of 560. Student gender proportions across the countries were all very close to 50%.

Table 3 displays the intraclass correlations (ICC), standard errors of ICCs, and the R^2 values. A small ICC indicated a smaller variability across schools within a country.

Examination of the table shows that DEU and NLD had the highest ICCs whereas DNK, NOR, and CNL had the lowest ICCs. The amount of variance explained by Model 2 at level 1 was nearly negligible across the countries. Model 1 and Model 3 were similar to each other in explaining the variances at level 1. However, all three models had unique (country-dependent) contributions in explaining the variances at

Table 2 Descriptive statistics for 2013 ICILS

Country	N	J	CIL	CIL:SD	AGE	MALE%	S_INTRST	S_BASEFF	C_RATSTD	C_ICTRES	T_USETCH
ABA	1076	68	458	87	44	46	51	51	29.9	50	50
AUS	5326	311	540	74	42	50	49	52	2.2	63	51
CHE	3225	98	514	64	44	52	48	49	10.0	50	50
CHL	3180	174	495	85	42	51	55	53	19.2	47	50
CNL	1556	118	515	78	40	47	53	51	4.7	55	50
COT	3377	193	535	70	40	50	51	51	7.4	58	51
CZE	3066	170	560	59	44	49	50	51	10.1	49	50
DEU	2225	136	529	70	46	51	48	50	11.5	52	50
DNK	1767	103	542	65	44	51	49	52	4.3	59	50
HKG	2089	118	515	89	39	53	50	48	9.0	59	50
HRV	2850	170	516	77	44	51	53	53	25.1	47	50
KOR	2888	150	536	85	43	51	46	49	19.7	51	50
LTU	2756	162	494	80	47	51	50	49	13.0	54	50
NLD	2197	121	536	78	43	52	46	52	5.3	61	50
NOR	2436	138	538	67	44	50	50	51	2.4	54	50
POL	2870	156	532	75	43	52	51	53	9.7	51	50
RUS	3626	206	518	72	46	50	48	51	17.4	52	50
SVK	2994	167	519	89	43	51	48	51	9.6	50	50
SVN	3740	218	510	65	44	52	50	53	14.4	55	50
THA	3646	198	385	97	40	50	50	40	14.2	48	50
TUR	2540	141	362	96	35	52	52	44	78.9	37	50

N: Total number of students in a country, J: Number of schools participated from a country, CIL: Mean computer and information literacy in a country, CIL:SD: Standard deviation of computer and information literacy in a country AGE: Mean age of teachers in a country, MALES%: Percentage of male students in a country, S_INTRST: Mean interest and enjoyment in using ICT, S_BASEFF: Mean of ICT self-efficacy or self-efficacy for basic/general skills, C_RATSTD: Mean ratio of school size and number of computers available for students in a country, C_ICTRES: Availability of ICT resources at school within a country, T_USETCH: Average use of ICT for teaching at school in a country

level 2. In addition, all three models explained more variance at level 2 than they did at level 1 in the majority of the countries.

Table 4 shows the descriptive statistics of the 2018 cycle of the ICILS study, including computer and information literacy (CIL) and computational thinking (CT) skills. Five countries (CHL, ITA, KAZ, RMO, and URY) did not collect data on CT, so that their CT scores were not available. KAZ and URY had the lowest CIL scores among the 14 countries. LUX and PRT had the lowest level of CT among the 9 countries that had data on CT.

Table 5 displays ICCs, and R^2 s for CIL and CT across the countries that participated in 2018. ICC for CIL was the lowest for KOR (0.14) and highest for DEU (0.58). Model 1 and Model 3 explained similar amounts of variances at level 1 for almost all of the countries. All three models explained large variances at level 2 for CIL at varying degrees. With regard to CIL that Model 1 explained more

Table 3 ICC estimates and proportion of additional variances explained (ICILS 2013)

Model	Null	Null	M1	M2	M3	M1	M2	M3
Statistic	ICC	ICC SE	R^2 at Level 1			R^2 at Level 2		
ABA	0.41	0.04	0.11	0.01	0.11	0.68	0.49	0.72
AUS	0.30	0.02	0.18	0.00	0.18	0.39	0.19	0.50
CHE	0.23	0.03	0.16	0.11	0.16	0.42	0.61	0.63
CHL	0.41	0.03	0.13	0.00	0.13	0.41	0.15	0.50
CNL	0.14	0.02	0.14	0.00	0.13	0.16	0.06	0.22
COT	0.16	0.01	0.15	0.00	0.15	0.26	0.02	0.27
CZE	0.27	0.02	0.09	0.00	0.09	0.17	0.09	0.23
DEU	0.59	0.03	0.09	0.00	0.09	0.14	0.20	0.30
DNK	0.13	0.02	0.12	0.00	0.12	0.22	0.19	0.34
HKG	0.55	0.03	0.12	0.00	0.12	0.17	0.10	0.25
HRV	0.18	0.02	0.17	0.00	0.17	0.31	0.16	0.39
KOR	0.20	0.02	0.22	0.00	0.22	0.21	0.02	0.23
LTU	0.42	0.03	0.16	0.00	0.16	0.32	0.19	0.42
NLD	0.56	0.03	0.13	0.00	0.13	0.22	0.56	0.64
NOR	0.14	0.01	0.15	0.00	0.15	0.11	0.21	0.29
POL	0.23	0.02	0.15	0.00	0.15	0.42	0.04	0.43
RUS	0.38	0.02	0.11	0.00	0.11	0.15	0.14	0.24
SVK	0.39	0.03	0.17	0.00	0.17	0.35	0.07	0.39
SVN	0.15	0.01	0.17	0.00	0.17	0.15	0.04	0.18
THA	0.47	0.03	0.07	0.00	0.07	0.16	0.18	0.31
TUR	0.42	0.03	0.12	0.00	0.12	0.24	0.36	0.46

ICC: Intra-class correlation coefficient, ICC SE: standard error estimate for ICC, Null: Empty model (model that does not include any covariate), M1: Model 1 (only student-level covariates), M2: Model 2 (only school/teacher level covariates), M3: Model 3 (both student and school/teacher level covariates). Sample sizes to run the models could be lower than the N displayed in Table 2 due to listwise deletion of cases having missing data

Table 4 Descriptive statistics for CIL and CT in ICILS 2018

Country	N	IIIIJK	CIL	k CIL.SD	CT	CT.SD	AGE o	MALE%	S_GENEFF	C_RATSTD	C_ICTRES	T_CLASACT
CHL	3092	178	482	85	–	–	41	49	52	17.7	45	49
DEU	3655	209	519	73	488	97	45	51	49	11.6	42	45
DNK	2404	143	553	64	528	81	45	51	52	5.9	57	58
DNW	1991	109	516	74	487	93	45	51	48	13.1	41	44
FIN	2546	144	527	79	504	95	46	51	52	3.2	59	48
FRA	2940	156	496	78	501	89	43	50	49	7.4	50	46
ITA	2810	150	461	78	–	–	52	52	50	14.9	47	47
KAZ	3371	183	392	101	–	–	42	52	46	22.9	52	54
KOR	2875	150	539	92	532	107	44	52	49	13.2	54	47
LUX	5401	38	482	80	460	103	41	53	48	3.7	51	47
PRT	3221	200	513	68	481	76	48	51	53	15.8	49	47
RMO	2852	150	553	64	–	–	46	51	51	13.7	53	55
URY	2613	166	443	101	–	–	41	50	50	23.6	55	50
USA	6790	263	517	77	496	103	42	51	51	1.6	59	49

variance at level 2 than Model 2 did for the majority of the countries, even though Model 1 had only level 1 predictors. Examination of ICCs for CT indicated that CT scores varied across schools as ICCs ranged from 0.11 to 0.49. This means that the variations across schools were country-specific. Model 1 and Model 3 explained more variance at level 1 than Model 2 did. Meantime, all three models explained significant amount of variations at level 2. Additionally, the ICCs for CT scores are not free from the ICCs of the CIL scores.

Given the results displayed in Tables 2, 3, 4, 5, MDES values were calculated for Model 1-3, and results were displayed in Figs. 2, 3, 4. First of all, countries differed in terms of MDES even if the same model and research design specifications were used. For instance, MDES was 1.7 for NLD and 0.8 for SVN when Model 1 was used, assuming 20 students from each of 10 schools participated in a study. Second, increasing number of schools (J) from 10 to 40 decreased the MDES by about 50% for Model 1 and around 65% for Model 2 and Model 3. However, the increasing number of students (N) within a school while keeping the number of schools (J) fixed can hardly decrease the MDES. The differences were small enough to be ignored. These findings are evenly held for the 2018 dataset as well; please see Fig. 3 for more details. Further, there were four countries that participated in both cycles: CHL, DEU, DNK, and KOR. CHL and KOR had higher MDES in 2013 compared to 2018 estimates; in contrast, DEU and DNK had MDES values that were stable from 2013 to 2018. As a result, it is hard to say whether the MDES estimates are stable overtime or not. Lastly, the MDES estimates for the countries with high ICC (e.g., DEU, HKG) were higher even if large sample sizes are selected compared to the countries with smaller ICCs (e.g., DNK, NOR).

MDES values with regard to CT scores for Model 1, Model 2, and Model 3 were displayed in Fig. 4. It was clear to see that countries differed from one another even if the same constraints were used to calculate MDES. Model 2 had slightly higher MDES values than Model 1 across the countries listed. Comparing Model 3 to the other two models, Model 3 was associated with smaller MDES but only for the conditions when the number of schools (J) were 20 or higher.

6 Conclusions and recommendations

This study is proposed to improve the design of CRTs in CIL and CT areas by estimating (i) intraclass correlation coefficients, (ii) the amount of variance explained given the covariates, and (iii) the MDES. The objective was accomplished by using the data provided by IEA. The findings showed that countries vary from one another in terms of the value of the ICCs, with differences of 0.44 in 2013 and 0.46 in 2018 for CIL scores. Meanwhile, the ICCs for the CT scores had a difference of 0.38 in 2018. Thus, ICCs were country-specific. Each country has its own degrees of variability among schools in terms of student CIL and CT scores due to its own background and culture. As a result, to ensure the design of cluster-randomized studies in CIL and CT field was neither underpowered (inaccurate) nor overpowered, which means sample resources are wasted, baseline data is necessary and essential. Our

Table 5 ICCs, R²s for CIL and CT in ICILS 2018

Model	Computer and information literacy						Computational thinking							
	Null	Null	M1	M2	M3	M3	Null	Null	M1	M2	M3	M1	M2	M3
Country	ICC	ICC SE	R ² at Level 1	R ² at Level 2	R ² at Level 2	R ² at Level 2	ICC	SE	R ² at Level 1	R ² at Level 1	R ² at Level 1	R ² at Level 2	R ² at Level 2	R ² at Level 2
CHL	0.32	0.02	0.15	0.03	0.15	0.47	—	—	—	—	—	—	—	—
DEU	0.58	0.02	0.13	0.05	0.13	0.52	0.49	0.02	0.11	0.07	0.11	0.00	0.29	0.34
DNK	0.19	0.02	0.20	0.05	0.20	0.40	0.11	0.01	0.17	0.04	0.17	0.32	0.03	0.33
DNW	0.40	0.03	0.08	0.00	0.08	0.37	0.36	0.03	0.06	0.00	0.06	0.03	0.26	0.27
FIN	0.17	0.02	0.26	0.10	0.26	0.40	0.16	0.02	0.20	0.09	0.20	0.32	0.18	0.38
FRA	0.22	0.02	0.21	0.07	0.21	0.46	0.22	0.02	0.15	0.06	0.15	0.46	0.43	0.52
ITA	0.28	0.02	0.22	0.02	0.22	0.47	—	—	—	—	—	—	—	—
KAZ	0.53	0.03	0.12	0.03	0.12	0.40	—	—	—	—	—	—	—	—
KOR	0.14	0.01	0.23	0.01	0.23	0.55	0.15	0.01	0.18	0.02	0.18	0.36	0.16	0.39
LUX	0.22	0.04	0.09	0.00	0.09	0.77	0.23	0.04	0.05	0.00	0.05	0.63	0.67	0.77
PRT	0.37	0.02	0.11	0.02	0.11	0.47	0.33	0.02	0.12	0.02	0.12	0.28	0.21	0.39
RMO	0.15	0.01	0.13	0.00	0.13	0.38	—	—	—	—	—	—	—	—
URY	0.47	0.03	0.25	0.10	0.25	0.56	—	—	—	—	—	—	—	—
USA	0.23	0.02	0.20	0.03	0.20	0.64	0.20	0.01	0.13	0.02	0.13	0.50	0.32	0.61

Fig. 2 2013 MDES estimates for CIL scores. Note: We used heat maps to make it easier for visual understanding; interested readers are advised to check the actual tables in Appendix 1

Fig. 3 2018 MDES estimates for CIL scores

Fig. 4 2018 MDES estimates for CT scores

findings provided such baseline standards for further researchers to explore in these areas.

The finding of ICCs in the United States is that the ICCs of CIL and CT are 0.23 and 0.20, respectively, which are similar to the means of the ICCs in academic outcomes that include reading, math, and science achievements. It indicates that the studies in CIL and CT fields in the U.S. may adopt a similar standard as their academic outcome studies. However, prior research suggested that the estimated value of ICC is context-specific, and the variability may occur when the participants are in different grades, school districts, or states. By far, there is not sufficient research in the field of CIL and CT in different situations yet. Therefore, we recommend conducting more empirical work that under different situations across the states to obtain a more accurate range of ICCs in CIL and CT.

Another finding is regarding R^2 . That is, the covariates at level 1 and level 2 explained different amounts the within/between-school variance on CIL and CT students' outcomes in many countries. However, the between-school variance can mostly be explained by level 1 covariates at level 2, rather than the level 2 covariates only. The student covariates at level 1 primarily explained the variances within the school. For example, the amounts of variance explained by covariates at level 1 (student level) in model 1 in the U.S. were less than 20%; in contrast, as we aggregated the same covariates to level 2, the amounts of explained variance increased to 50%. In addition, the largest amounts of explained between-school variance were in model 3 (with both level 1 and level 2 covariates). As mentioned above, the percent of variance explained by covariates in model 3 in the U.S. were up to 64% in CIL and 61% in CT, which were 10% higher than the amounts that only include aggregated student-level covariates at teacher/school level. However, the influence of covariates on reducing between-school variance depends strongly on the country (Kelcey et al., 2016; Spybrook et al., 2016). For example, the between-school variances explained by the teacher/school level covariates were up to 0.77 in LUX in CIL and CT, while it was 0.37 in DNW in CIL and 0.27 in CT.

This leads to the second key finding related to R^2 . The explanatory power of covariates varies by country. This is evidenced in the higher percentage of the achievement differences explained by level 2 covariates in LUX (77% in CIL and CT) compared to DNW (37% in CIL and 27% in CT). This indicates that the current covariates used in the analyses could not sufficiently reduce the school variance in DNW. Therefore, we encourage further research to explore other typical covariates locally to better explain the variability across the schools in different countries.

The third conclusion derived from our results is regarding MDES. As we can see from the formulas, the ICCs and MDES are related; that is, the smaller the ICC value, the smaller the MDES. The R^2 and MDES are also related; the larger the R^2 , the smaller the MDES. The MDES calculations in this work were based on the empirical estimates of the ICCs and R^2 's as well as a fixed total number of clusters, cluster size, a two-tailed hypothesis test, a significance level set to 0.05, and a power set to 0.80. Findings suggested that just like the ICC and the R^2 s, the MDES is context-specific as well. For example, as we can see in Figs. 3 and 4, the ICCs of CIL and CT in the U.S. and DNW were 0.36 and 0.2, respectively. Consequently, the MDES were 1.4 and 0.7 in these two different

countries. It suggested that larger ICC results in a larger MDES, which causes less accurate conclusions. In addition, the MDES was reduced by adding on the covariates. The more percent of variance explained, the smaller MDES were observed. Therefore, given the different empirical estimates of the ICCs and R^2 , the MDES is context-specific.

Lastly, the MDES for CIL and CT decreased as the number of schools increased rather than the number of students inflated. As seen in Figs. 3 and 4, for instance, in the U.S., increasing the number of students within a school from 20 to 60 to participate in a CRT is less likely to reduce the MDES to a significant degree. However, increasing the number of schools from 10 to 40 while keeping the sample size in each school fixed makes a huge decrease in MDES (less than %50 for the majority of instances). This conclusion holds even for the countries with very low ICCs, such as SVN, NOR, and DNK. As a result, we recommend that when planning sample size for CRTs, increasing the number of schools would help achieve smaller MDES values. In designing studies that are effective, researchers always struggle when planning the sampling stage. Small samples may lead to type II errors, while too large sample could waste economic resources. With this study, we encourage the use of CRTs when planning interventions for CIL or CT (Appendix Tables 6, 7, 8, 9).

7 Limitations

First of all, we incorporated teacher covariates into the analysis at the school level by using a central tendency quantity (e.g., mean, median) as if they were school characteristics. The reason for averaging teacher data was that there was no linkage between teacher and student data due to random sampling of teachers from a particular school. The only way to keep the teacher data in the analysis was to average them at the school level since the dataset had school ID for teachers; thus, it would have been a better strategy if we were able to incorporate teacher data at level 2 and school data at level 3.

Second, the data collection procedures were not exactly identical for the two cycles, which forced us to make some sort of subjective decisions when we choose covariates. The reason is that the survey was modified from 2013 to 2018 for some variables. For instance, one of the covariates at the teacher level is ‘*Use of ICT for teaching at school*’ (T_USETCH) in 2013, while the other one is ‘*Use of ICT for classroom activities*’ (T_CLASACT) in 2018. There is another example regarding student experience with computers. For the 2013 cycle, we used the ‘*Interest and enjoyment in using ICT*’ (S_INTRST); in contrast, ‘*Students’ computer experience in years*’ (S_EXCOMP) was used in 2018. We showed the differences in data collection procedures in Table 1.

Lastly, the number of countries that participated in both cycles is only four (CHL, DNK, DEU, and KOR); therefore, it is hard to make general conclusions when ICC, R^2 , and MDES of CIL scores from 2013 to 2018 were compared.

Appendix 1. MDES Values Displayed in Figs. 2, 3, and 4

Table 6 2013 and 2018 CIL MDES for Model 1 (student level covariates)

	N	20	20	20	20	40	40	40	40	60	60	60	60
Year	J	10	20	30	40	10	20	30	40	10	20	30	40
2013	ABA	1.34	0.88	0.70	0.60	1.32	0.86	0.69	0.59	1.31	0.86	0.69	0.59
	AUS	1.20	0.79	0.63	0.54	1.18	0.78	0.62	0.53	1.18	0.77	0.62	0.53
	CHE	1.04	0.68	0.55	0.47	1.01	0.66	0.53	0.45	1.00	0.66	0.52	0.45
	CHL	1.34	0.88	0.70	0.60	1.32	0.87	0.69	0.59	1.31	0.86	0.69	0.59
	CNL	0.92	0.60	0.48	0.41	0.88	0.57	0.46	0.39	0.86	0.57	0.45	0.39
	COT	0.96	0.63	0.50	0.43	0.92	0.60	0.48	0.41	0.91	0.60	0.48	0.41
	CZE	1.10	0.72	0.58	0.49	1.07	0.70	0.56	0.48	1.06	0.69	0.55	0.48
	DEU	1.60	1.05	0.84	0.72	1.59	1.04	0.83	0.71	1.58	1.04	0.83	0.71
	DNK	0.97	0.64	0.51	0.44	0.93	0.61	0.49	0.42	0.92	0.60	0.48	0.41
	HKG	1.50	0.98	0.79	0.67	1.49	0.97	0.78	0.67	1.48	0.97	0.78	0.67
	HRV	0.93	0.61	0.49	0.42	0.90	0.59	0.47	0.40	0.89	0.58	0.46	0.40
	KOR	0.99	0.65	0.52	0.45	0.96	0.63	0.51	0.43	0.95	0.62	0.50	0.43
	LTU	1.31	0.86	0.69	0.59	1.30	0.85	0.68	0.58	1.29	0.85	0.68	0.58
	NLD	1.70	1.11	0.89	0.76	1.69	1.11	0.89	0.76	1.69	1.11	0.88	0.76
	NOR	0.88	0.58	0.46	0.40	0.84	0.55	0.44	0.38	0.83	0.54	0.43	0.37
	POL	1.03	0.67	0.54	0.46	0.99	0.65	0.52	0.45	0.98	0.64	0.52	0.44
	RUS	1.28	0.84	0.67	0.57	1.25	0.82	0.66	0.56	1.25	0.82	0.65	0.56
	SVK	1.30	0.85	0.68	0.59	1.28	0.84	0.67	0.58	1.28	0.84	0.67	0.58
	SVN	0.85	0.56	0.45	0.38	0.81	0.53	0.42	0.36	0.79	0.52	0.42	0.36
	THA	1.36	0.89	0.71	0.61	1.34	0.88	0.70	0.60	1.33	0.87	0.70	0.60
	TUR	1.45	0.95	0.76	0.65	1.43	0.94	0.75	0.64	1.43	0.94	0.75	0.64
2018	CHL	1.20	0.79	0.63	0.54	1.17	0.77	0.61	0.53	1.16	0.76	0.61	0.52
	DEU	1.66	1.09	0.87	0.75	1.65	1.08	0.86	0.74	1.65	1.08	0.86	0.74
	DNK	0.96	0.63	0.51	0.43	0.93	0.61	0.49	0.42	0.92	0.60	0.48	0.41
	DNW	1.33	0.87	0.70	0.60	1.30	0.86	0.68	0.59	1.30	0.85	0.68	0.58
	FIN	0.92	0.60	0.48	0.41	0.88	0.58	0.46	0.40	0.86	0.57	0.45	0.39
	FRA	0.96	0.63	0.50	0.43	0.93	0.61	0.49	0.42	0.91	0.60	0.48	0.41
	ITA	1.07	0.70	0.56	0.48	1.05	0.69	0.55	0.47	1.04	0.68	0.54	0.47
	KAZ	1.50	0.98	0.79	0.67	1.48	0.97	0.78	0.67	1.48	0.97	0.78	0.67
	KOR	0.81	0.53	0.43	0.37	0.77	0.50	0.40	0.35	0.75	0.49	0.40	0.34
	LUX	0.85	0.56	0.45	0.38	0.81	0.53	0.42	0.36	0.79	0.52	0.42	0.36
	PRT	1.21	0.80	0.64	0.55	1.19	0.78	0.62	0.53	1.18	0.77	0.62	0.53
	RMO	0.85	0.55	0.44	0.38	0.80	0.52	0.42	0.36	0.78	0.51	0.41	0.35
	URY	1.38	0.91	0.72	0.62	1.36	0.89	0.72	0.61	1.36	0.89	0.71	0.61
	USA	0.98	0.64	0.51	0.44	0.95	0.62	0.50	0.43	0.93	0.61	0.49	0.42

Table 7 2013 and 2018 CIL MDES for Model 2 (school level covariates)

	N	20	20	20	20	40	40	40	40	60	60	60	60
Year	J	10	20	30	40	10	20	30	40	10	20	30	40
2013	ABA	1.29	0.67	0.53	0.45	1.25	0.65	0.51	0.44	1.24	0.64	0.50	0.43
	AUS	1.26	0.65	0.51	0.44	1.22	0.63	0.49	0.42	1.20	0.62	0.49	0.42
	CHE	0.94	0.49	0.38	0.33	0.87	0.45	0.35	0.30	0.84	0.44	0.34	0.29
	CHL	1.55	0.80	0.63	0.54	1.52	0.78	0.62	0.53	1.50	0.78	0.61	0.52
	CNL	1.01	0.52	0.41	0.35	0.94	0.48	0.38	0.32	0.91	0.47	0.37	0.32
	COT	1.09	0.56	0.44	0.38	1.03	0.53	0.42	0.36	1.00	0.52	0.41	0.35
	CZE	1.37	0.70	0.56	0.47	1.32	0.68	0.54	0.46	1.30	0.67	0.53	0.45
	DEU	1.73	0.89	0.70	0.60	1.71	0.88	0.70	0.59	1.71	0.88	0.69	0.59
	DNK	0.95	0.49	0.39	0.33	0.88	0.45	0.36	0.30	0.85	0.44	0.34	0.29
	HKG	1.76	0.91	0.72	0.61	1.74	0.90	0.71	0.60	1.73	0.89	0.70	0.60
	HRV	1.15	0.59	0.47	0.40	1.09	0.56	0.44	0.38	1.07	0.55	0.43	0.37
	KOR	1.25	0.64	0.51	0.43	1.19	0.61	0.48	0.41	1.17	0.60	0.48	0.41
	LTU	1.55	0.80	0.63	0.54	1.51	0.78	0.62	0.53	1.50	0.77	0.61	0.52
	NLD	1.15	0.59	0.47	0.40	1.13	0.58	0.46	0.39	1.12	0.58	0.45	0.39
	NOR	1.04	0.53	0.42	0.36	0.96	0.50	0.39	0.33	0.94	0.48	0.38	0.32
	POL	1.31	0.68	0.53	0.46	1.26	0.65	0.51	0.44	1.25	0.64	0.51	0.43
	RUS	1.55	0.80	0.63	0.54	1.52	0.78	0.62	0.53	1.51	0.78	0.61	0.52
	SVK	1.53	0.79	0.62	0.53	1.50	0.77	0.61	0.52	1.48	0.76	0.60	0.51
	SVN	1.11	0.57	0.45	0.38	1.04	0.53	0.42	0.36	1.01	0.52	0.41	0.35
	THA	1.70	0.88	0.69	0.59	1.67	0.86	0.68	0.58	1.66	0.86	0.67	0.58
	TUR	1.29	0.66	0.52	0.45	1.25	0.65	0.51	0.43	1.24	0.64	0.50	0.43
	CHL	1.55	0.80	0.63	0.54	1.52	0.78	0.62	0.53	1.50	0.78	0.61	0.52
	CHL	1.48	0.76	0.60	0.51	1.44	0.74	0.58	0.50	1.42	0.73	0.58	0.49
	DEU	1.73	0.89	0.70	0.60	1.71	0.88	0.70	0.59	1.71	0.88	0.69	0.59
	DEU	1.45	0.75	0.59	0.50	1.43	0.74	0.58	0.50	1.42	0.73	0.58	0.49
	DNK	0.95	0.49	0.39	0.33	0.88	0.45	0.36	0.30	0.85	0.44	0.34	0.29
	DNK	1.25	0.64	0.51	0.43	1.19	0.61	0.48	0.41	1.17	0.60	0.47	0.41
	DNW	1.48	0.76	0.60	0.51	1.44	0.74	0.58	0.50	1.43	0.74	0.58	0.50
	FIN	1.16	0.60	0.47	0.40	1.10	0.57	0.45	0.38	1.08	0.55	0.44	0.37
	FRA	1.20	0.62	0.49	0.42	1.14	0.59	0.46	0.40	1.12	0.58	0.46	0.39
2018	ITA	1.38	0.71	0.56	0.48	1.33	0.69	0.54	0.46	1.32	0.68	0.53	0.46
	KAZ	1.72	0.89	0.70	0.60	1.69	0.87	0.69	0.59	1.69	0.87	0.68	0.59
	KOR	1.25	0.64	0.51	0.43	1.19	0.61	0.48	0.41	1.17	0.60	0.48	0.41
	KOR	1.08	0.55	0.44	0.37	1.00	0.52	0.41	0.35	0.98	0.50	0.40	0.34
	LUX	0.94	0.48	0.38	0.33	0.86	0.44	0.35	0.30	0.83	0.43	0.34	0.29
	PRT	1.47	0.76	0.60	0.51	1.43	0.74	0.58	0.50	1.42	0.73	0.58	0.49
	RMO	1.08	0.55	0.44	0.37	1.01	0.52	0.41	0.35	0.98	0.51	0.40	0.34
	URY	1.68	0.87	0.68	0.58	1.65	0.85	0.67	0.57	1.64	0.85	0.67	0.57
	USA	1.15	0.59	0.47	0.40	1.09	0.56	0.44	0.38	1.07	0.55	0.44	0.37

Table 8 2013 and 2018 CIL MDES for Model 3 (student and school level covariates)

Year	N	20	20	20	20	40	40	40	40	60	60	60	60
	J	10	20	30	40	10	20	30	40	10	20	30	40
2013	ABA	1.02	0.53	0.42	0.36	0.98	0.50	0.40	0.34	0.96	0.50	0.39	0.33
	AUS	1.01	0.52	0.41	0.35	0.96	0.49	0.39	0.33	0.94	0.49	0.38	0.33
	CHE	0.90	0.47	0.37	0.31	0.84	0.43	0.34	0.29	0.82	0.42	0.33	0.28
	CHL	1.17	0.60	0.47	0.41	1.13	0.58	0.46	0.39	1.12	0.58	0.45	0.39
	CNL	0.84	0.43	0.34	0.29	0.77	0.39	0.31	0.27	0.74	0.38	0.30	0.26
	COT	0.92	0.47	0.37	0.32	0.86	0.44	0.35	0.30	0.83	0.43	0.34	0.29
	CZE	1.18	0.61	0.48	0.41	1.13	0.58	0.46	0.39	1.11	0.57	0.45	0.39
	DEU	1.64	0.84	0.67	0.57	1.62	0.84	0.66	0.56	1.62	0.83	0.66	0.56
	DNK	0.82	0.42	0.33	0.29	0.75	0.38	0.30	0.26	0.72	0.37	0.29	0.25
	HKG	1.65	0.85	0.67	0.57	1.63	0.84	0.66	0.57	1.62	0.84	0.66	0.56
	HRV	1.00	0.51	0.41	0.35	0.94	0.48	0.38	0.33	0.92	0.47	0.37	0.32
	KOR	1.08	0.56	0.44	0.37	1.03	0.53	0.42	0.36	1.01	0.52	0.41	0.35
	LTU	1.30	0.67	0.53	0.45	1.27	0.66	0.52	0.44	1.26	0.65	0.51	0.44
	NLD	1.11	0.57	0.45	0.38	1.08	0.56	0.44	0.38	1.08	0.55	0.44	0.37
	NOR	0.94	0.48	0.38	0.33	0.87	0.45	0.35	0.30	0.85	0.44	0.35	0.30
	POL	0.96	0.49	0.39	0.33	0.90	0.46	0.36	0.31	0.88	0.45	0.36	0.30
	RUS	1.47	0.76	0.60	0.51	1.43	0.74	0.58	0.50	1.42	0.73	0.58	0.49
	SVK	1.26	0.65	0.51	0.44	1.22	0.63	0.50	0.42	1.21	0.62	0.49	0.42
	SVN	1.01	0.52	0.41	0.35	0.95	0.49	0.38	0.33	0.93	0.48	0.38	0.32
	THA	1.51	0.78	0.61	0.53	1.48	0.76	0.60	0.52	1.47	0.76	0.60	0.51
	TUR	1.18	0.61	0.48	0.41	1.14	0.59	0.46	0.40	1.13	0.58	0.46	0.39
2018	CHL	1.13	0.58	0.46	0.39	1.09	0.56	0.44	0.38	1.07	0.55	0.43	0.37
	DEU	1.24	0.64	0.51	0.43	1.22	0.63	0.50	0.42	1.22	0.63	0.49	0.42
	DNK	0.97	0.50	0.40	0.34	0.91	0.47	0.37	0.32	0.89	0.46	0.36	0.31
	DNW	1.32	0.68	0.54	0.46	1.29	0.66	0.52	0.45	1.28	0.66	0.52	0.44
	FIN	0.95	0.49	0.39	0.33	0.89	0.46	0.36	0.31	0.87	0.45	0.35	0.30
	FRA	1.01	0.52	0.41	0.35	0.95	0.49	0.39	0.33	0.93	0.48	0.38	0.32
	ITA	1.10	0.57	0.45	0.38	1.05	0.54	0.43	0.37	1.04	0.53	0.42	0.36
	KAZ	1.53	0.79	0.62	0.53	1.51	0.78	0.61	0.52	1.50	0.77	0.61	0.52
	KOR	0.81	0.42	0.33	0.28	0.74	0.38	0.30	0.26	0.71	0.37	0.29	0.25
	LUX	0.70	0.36	0.28	0.24	0.60	0.31	0.24	0.21	0.56	0.29	0.23	0.19
	PRT	1.26	0.65	0.51	0.44	1.21	0.62	0.49	0.42	1.20	0.62	0.49	0.42
	RMO	0.92	0.48	0.38	0.32	0.85	0.44	0.34	0.29	0.82	0.42	0.33	0.29
	URY	1.23	0.63	0.50	0.43	1.19	0.62	0.49	0.41	1.18	0.61	0.48	0.41
	USA	0.84	0.43	0.34	0.29	0.77	0.40	0.31	0.27	0.75	0.39	0.30	0.26

Table 9 2018 CT MDES estimates for M1, M2 and M3

	J	10	20	30	40	10	20	30	40	10	20	30	40
	N	20	20	20	20	40	40	40	40	60	60	60	60
M1	DEU	1.45	0.95	0.76	0.65	1.43	0.94	0.75	0.64	1.43	0.94	0.75	0.64
	DNK	0.79	0.52	0.41	0.35	0.74	0.48	0.39	0.33	0.72	0.47	0.38	0.32
	DNW	1.26	0.83	0.66	0.57	1.24	0.81	0.65	0.56	1.23	0.81	0.65	0.55
	FIN	0.88	0.58	0.46	0.40	0.84	0.55	0.44	0.38	0.83	0.54	0.44	0.37
	FRA	1.00	0.66	0.53	0.45	0.97	0.64	0.51	0.44	0.96	0.63	0.50	0.43
	KOR	0.86	0.56	0.45	0.39	0.82	0.54	0.43	0.37	0.80	0.53	0.42	0.36
	LUX	1.05	0.69	0.55	0.47	1.01	0.66	0.53	0.45	1.00	0.65	0.52	0.45
	PRT	1.21	0.80	0.64	0.55	1.19	0.78	0.62	0.53	1.18	0.77	0.62	0.53
	USA	0.98	0.64	0.52	0.44	0.95	0.62	0.50	0.43	0.93	0.61	0.49	0.42
	DEU	1.6	0.83	0.66	0.56	1.59	0.82	0.64	0.55	1.58	0.81	0.64	0.55
M2	DNK	1.0	0.53	0.42	0.36	0.96	0.49	0.39	0.33	0.93	0.48	0.38	0.32
	DNW	1.5	0.75	0.59	0.50	1.41	0.73	0.57	0.49	1.40	0.72	0.57	0.49
	FIN	1.1	0.56	0.44	0.38	1.02	0.53	0.41	0.35	1.00	0.51	0.40	0.35
	FRA	1.1	0.54	0.43	0.37	0.99	0.51	0.40	0.34	0.97	0.50	0.39	0.34
	KOR	1.1	0.55	0.43	0.37	0.99	0.51	0.40	0.34	0.96	0.50	0.39	0.33
	LUX	0.9	0.46	0.37	0.31	0.83	0.43	0.34	0.29	0.80	0.41	0.32	0.28
	PRT	1.4	0.74	0.58	0.50	1.39	0.71	0.56	0.48	1.37	0.71	0.56	0.48
	USA	1.1	0.57	0.45	0.38	1.04	0.53	0.42	0.36	1.02	0.52	0.41	0.35
	DEU	1.44	0.74	0.59	0.50	1.42	0.73	0.57	0.49	1.41	0.72	0.57	0.49
	DNK	0.86	0.44	0.35	0.30	0.79	0.41	0.32	0.27	0.76	0.39	0.31	0.26
M3	DNW	1.36	0.70	0.55	0.47	1.32	0.68	0.54	0.46	1.31	0.67	0.53	0.45
	FIN	0.91	0.47	0.37	0.31	0.84	0.43	0.34	0.29	0.82	0.42	0.33	0.28
	FRA	0.90	0.47	0.37	0.31	0.84	0.43	0.34	0.29	0.82	0.42	0.33	0.28
	KOR	0.91	0.47	0.37	0.32	0.85	0.44	0.34	0.29	0.82	0.42	0.33	0.29
	LUX	0.70	0.36	0.28	0.24	0.60	0.31	0.24	0.21	0.57	0.29	0.23	0.20
	PRT	1.21	0.62	0.49	0.42	1.17	0.60	0.47	0.40	1.15	0.59	0.47	0.40
	USA	0.83	0.43	0.34	0.29	0.76	0.39	0.31	0.26	0.73	0.38	0.30	0.25

Data availability The data used in this study were from the 2013 and 2018 cycles of the International Computer and Information Literacy Study.

Code availability The R code used in the study was added as an appendix.

Declarations

Conflict of interest No conflict of interest was reported by the author(s).

Ethics approval Not applicable.

References

- Alkan, M., & Meinck, S. (2016). The relationship between students' use of ICT for social communication and their computer and information literacy. *Large-Scale Assessments in Education*, 4(1), 1–17. <https://doi.org/10.1186/s40536-016-0029-z>
- American Library Association. (2000). *Information literacy competency standards for higher education*. <https://repository.arizona.edu/handle/10150/105645>. Accessed 1 Mar 2021
- Aydin, M. (2021). Does the digital divide matter? Factors and conditions that promote ICT literacy. *Tele-matics and Informatics*, 58, 101536. <https://doi.org/10.1016/j.tele.2020.101536>
- Bailey, P., Kelley, C., Nguyen, T., Huo, H., Kjeldsen, C., & Bailey, M. P. (2020). *Package 'WeMix'*. <https://cran.r-project.org/web/packages/WeMix/index.html>. Accessed 5 Dec 2020
- Bawden, D. (2008). & others. (2008). Origins and concepts of digital literacy. *Digital literacies: Concepts. Policies and Practices*, 30, 17–32.
- Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review*, 19(5), 547–556. <https://doi.org/10.1177/0193841X9501900504>
- Bos, W., Eickelmann, B., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., Senkbeil, M., Schulz-Zander, R., & Wendt, H. (2014). *Computer-und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich*. Waxmann.
- Bulus, M., Dong, N., Kelcey, B., & Spybrook, J. (2019). PowerUPR: Power analysis tools for multilevel randomized experiments. *R Package Version*, 1.0.4. <https://cran.r-project.org/web/packages/PowerUPR/index.html>. Accessed 20 Nov 2020
- Cetin, I. (2016). Preservice teachers' introduction to computing: Exploring utilization of scratch. *Journal of Educational Computing Research*, 54(7), 997–1021. <https://doi.org/10.1177/0735633116642774>
- Ching, Y.-H., Hsu, Y.-C., & Baldwin, S. (2018). Developing computational thinking with educational technologies for young learners. *TechTrends*, 62(6), 563–573. <https://doi.org/10.1007/s11528-018-0292-7>
- Dede, C. (2010). Comparing frameworks for 21st century skills. In J. Bellanca & R. Brandt (Eds.), *21st century skills: Rethinking how students learn* (pp. 51–76). Solution Tree Press.
- Donner, A., & Koval, J. (1982). Design considerations in the estimation of intraclass correlation. *Annals of Human Genetics*, 46(3), 271–277. <https://doi.org/10.1111/j.1469-1809.1982.tb00718.x>
- Drossel, K., & Eickelmann, B. (2017). Teachers' participation in professional development concerning the implementation of new technologies in class: A latent class analysis of teachers and the relationship with the use of computers, ICT self-efficacy and emphasis on teaching ICT skills. *Large-Scale Assessments in Education*, 5(1), 1–13. <https://doi.org/10.1186/s40536-017-0053-7>
- Drossel, K., Eickelmann, B., & Gerick, J. (2017a). Predictors of teachers' use of ICT in school—the relevance of school characteristics, teachers' attitudes and teacher collaboration. *Education and Information Technologies*, 22(2), 551–573. <https://doi.org/10.1007/s10639-016-9476-y>
- Drossel, K., Eickelmann, B., & Schulz-Zander, R. (2017b). Determinants of teachers' collaborative use of information and communications technology for teaching and learning: A European perspective. *European Educational Research Journal*, 16(6), 781–799. <https://doi.org/10.1177/1474904116655811>
- Drossel, K., Eickelmann, B., & Vennemann, M. (2020). Schools overcoming the digital divide: In depth analyses towards organizational resilience in the computer and information literacy domain. *Large-Scale Assessments in Education*, 8(1), 1–19. <https://doi.org/10.1186/s40536-020-00087-w>
- Durak, H. Y., & Saritepeci, M. (2018). Analysis of the relation between computational thinking skills and various variables with the structural equation model. *Computers & Education*, 116, 191–202. <https://doi.org/10.1016/j.compedu.2017.09.004>
- Eickelmann, B., & Vennemann, M. (2017). Teachers' attitudes and beliefs regarding ICT in teaching and learning in European countries. *European Educational Research Journal*, 16(6), 733–761. <https://doi.org/10.1177/1474904117725899>
- Ferrari, A. (2012). *Digital competence in practice: An analysis of frameworks*. Joint Research Centre, Institute for Prospective Technological Studies. <https://doi.org/10.2791/82116>
- Fraillon, J., & Ainley, J. (2010). *The IEA international study of computer and information literacy (ICILS)*. Retrieved from <https://cms.education.gov.il/NR/rdonlyres/A07C37A9-8709-4B9B-85C0-63B9D57F1715/167512/ICILSProjectDescription1.pdf>. Accessed 25 Dec 2020

- 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 International Publishing. <https://doi.org/10.1007/978-3-319-14222-7>
- Fraillon, J., Ainley, J., Schulz, W., Duckworth, D., & Friedman, T. (2020). *Preparing for life in a digital age: IEA international computer and information literacy study 2018 international report*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-38781-5>
- Fraillon, J., Ainley, J., Schulz, W., Duckworth, D., & Friedman, T. (2019). *IEA international computer and information literacy study 2018 assessment framework*. Springer Nature.
- Garneli, V., & Chorianopoulos, K. (2018). Programming video games and simulations in science education: Exploring computational thinking through code analysis. *Interactive Learning Environments*, 26(3), 386–401. <https://doi.org/10.1080/10494820.2017.1337036>
- Garneli, V., & Chorianopoulos, K. (2019). The effects of video game making within science content on student computational thinking skills and performance. *Interactive Technology and Smart Education*, 16(4), 301–318. <https://doi.org/10.1108/ITSE-11-2018-0097>
- Gebhardt, E., Thomson, S., Ainley, J., & Hillman, K. (2019). *Gender differences in computer and information literacy: An in-depth analysis of data from ICILS*. Springer Nature.
- 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), 1–13. <https://doi.org/10.1186/s40536-017-0037-7>
- Guggemos, J. (2021). On the predictors of computational thinking and its growth at the high-school level. *Computers & Education*, 161, 1–15. <https://doi.org/10.1016/j.compedu.2020.104060>
- Hatlevik, O. E., Throndsen, I., Loi, M., & Gudmundsdottir, G. B. (2018). Students' ICT self-efficacy and computer and information literacy: Determinants and relationships. *Computers & Education*, 118, 107–119. <https://doi.org/10.1016/j.compedu.2017.11.011>
- Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, 29(1), 60–87. <https://doi.org/10.3102/0162373707299706>
- Hickmott, D., Prieto-Rodriguez, E., & Holmes, K. (2018). A scoping review of studies on computational thinking in K–12 mathematics classrooms. *Digital Experiences in Mathematics Education*, 4(1), 48–69. <https://doi.org/10.1007/s40751-017-0038-8>
- Hoffman, M., & Blake, J. (2003). Computer literacy: Today and tomorrow. *Journal of Computing Sciences in Colleges*, 18(5), 221–233. Retrieved from <https://dl.acm.org/doi/abs/10.5555/771832.771865>
- 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>
- Ilic, U., Haseski, H. İ., & Tugtekin, U. (2018). Publication trends over 10 years of computational thinking research. *Contemporary Educational Technology*, 9(2), 131–153. <https://doi.org/10.30935/cet.414798>
- Ioannou, A., & Makridou, E. (2018). Exploring the potentials of educational robotics in the development of computational thinking: A summary of current research and practical proposal for future work. *Education and Information Technologies*, 23(6), 2531–2544. <https://doi.org/10.1007/s10639-018-9729-z>
- ISTE. (2016). *Computational thinking competencies*. <https://www.iste.org/standards/computational-thinking>. Accessed 20 Mar 2021
- Jaipal-Jamani, K., & Angeli, C. (2017). Effect of robotics on elementary preservice teachers' self-efficacy, science learning, and computational thinking. *Journal of Science Education and Technology*, 26(2), 175–192. <https://doi.org/10.1007/s10956-016-9663-z>
- Jun, S., Han, S., & Kim, S. (2017). Effect of design-based learning on improving computational thinking. *Behaviour & Information Technology*, 36(1), 43–53. <https://doi.org/10.1080/0144929X.2016.1188415>
- Kalelioglu, F., Gulbahar, Y., & Kukul, V. (2016). A framework for computational thinking based on a systematic research review. *Baltic Journal of Modern Computing*, 4(3), 583–596. Retrieved from <https://www.bjmc.lu.lv/contents/vol-42016-no-3>. Accessed 20 Mar 2021
- Kelcey, B., & Phelps, G. (2013). Considerations for designing group randomized trials of professional development with teacher knowledge outcomes. *Educational Evaluation and Policy Analysis*, 35(3), 370–390. <https://doi.org/10.3102/0162373713482766>

- Kelcey, B., Shen, Z., & Spybrook, J. (2016). Intraclass correlation coefficients for designing cluster-randomized trials in sub-Saharan Africa education. *Evaluation Review*, 40(6), 500–525. <https://doi.org/10.1177/0193841X16660246>
- Kuo, W.-C., & Hsu, T.-C. (2020). Learning computational thinking without a computer: How computational participation happens in a computational thinking board game. *The Asia-Pacific Education Researcher*, 29(1), 67–83. <https://doi.org/10.1007/s40299-019-00479-9>
- Lau, W. W., & Yuen, A. H. (2014). Developing and validating of a perceived ICT literacy scale for junior secondary school students: Pedagogical and educational contributions. *Computers & Education*, 78, 1–9. <https://doi.org/10.1016/j.compedu.2014.04.016>
- Livingstone, S., & van der Graaf, S. (2008). Media Literacy. In: Donsbach, W. (Ed.). *International Encyclopedia of Communication*. Malden and Oxford, UK: Wiley-Blackwell. Retrieved from http://www.shenja.org/Shenja_van_der_Graaf/playlist_files/Media_Literacy_final08.pdf. Accessed 25 July 2021
- 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>
- Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education*, 59(3), 1065–1078. <https://doi.org/10.1016/j.compedu.2012.04.016>
- 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>
- Pala, F. K., & Mihçi Türker, P. (2021). The effects of different programming trainings on the computational thinking skills. *Interactive Learning Environments*, 29(7), 1090–1100. <https://doi.org/10.1080/10494820.2019.1635495>
- Punter, R. A., Meelissen, M. R., & Glas, C. A. (2017). Gender differences in computer and information literacy: An exploration of the performances of girls and boys in ICILS 2013. *European Educational Research Journal*, 16(6), 762–780. <https://doi.org/10.1177/1474904116672468>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>. Accessed 15 Sept 2020
- Rasch, G. (1993). *Probabilistic models for some intelligence and attainment tests*. ERIC.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Sage: Thousand Oaks, CA.
- Rohatgi, A., Scherer, R., & Hatlevik, O. E. (2016). The role of ICT self-efficacy for students' ICT use and their achievement in a computer and information literacy test. *Computers & Education*, 102, 103–116. <https://doi.org/10.1016/j.compedu.2016.08.001>
- 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>
- Siddiq, F., Scherer, R., & Tondeur, J. (2016). Teachers' emphasis on developing students' digital information and communication skills (TEDDICS): A new construct in 21st century education. *Computers & Education*, 92, 1–14. <https://doi.org/10.1016/j.compedu.2015.10.006>
- Spybrook, J., Westine, C. D., & Taylor, J. A. (2016). Design parameters for impact research in science education: A multistate analysis. *AERA Open*, 2(1), 1–15. <https://doi.org/10.1177/2332858415625975>
- Sung, W., Ahn, J., & Black, J. B. (2017). Introducing computational thinking to young learners: Practicing computational perspectives through embodiment in mathematics education. *Technology, Knowledge and Learning*, 22(3), 443–463. <https://doi.org/10.1007/s10758-017-9328-x>
- Tang, K.-Y., Chou, T.-L., & Tsai, C.-C. (2020). A content analysis of computational thinking research: An international publication trends and research typology. *The Asia-Pacific Education Researcher*, 29(1), 9–19. <https://doi.org/10.1007/s40299-019-00442-8>
- Tikva, C., & Tambouris, E. (2021). Mapping computational thinking through programming in K-12 education: A conceptual model based on a systematic literature review. *Computers & Education*, 162, 104083. <https://doi.org/10.1016/j.compedu.2020.104083>
- Tsai, C.-W., Shen, P.-D., Tsai, M.-C., & Chen, W.-Y. (2017). Exploring the effects of web-mediated computational thinking on developing students' computing skills in a ubiquitous learning environment. *Interactive Learning Environments*, 25(6), 762–777. <https://doi.org/10.1080/10494820.2016.1181093>
- van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577–588. <https://doi.org/10.1016/j.chb.2017.03.010>

- Varank, İ. (2006). A comparison of a computer-based and a lecture-based computer literacy course: A Turkish case. *Eurasia Journal of Mathematics, Science and Technology Education*, 2(3), 112–123. <https://doi.org/10.12973/ejmste/75469>
- Voogt, J., Fisser, P., Good, J., Mishra, P., & Yadav, A. (2015). Computational thinking in compulsory education: Towards an agenda for research and practice. *Education and Information Technologies*, 20(4), 715–728. <https://doi.org/10.1007/s10639-015-9412-6>
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33. <https://doi.org/10.1145/1118178.1118215>
- Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1881), 3717–3725. <https://doi.org/10.1098/rsta.2008.0118>
- Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational thinking in elementary and secondary teacher education. *ACM Transactions on Computing Education*, 14(1), 1–16. <https://doi.org/10.1145/2576872>
- Zhang, L., & Nouri, J. (2019). A systematic review of learning computational thinking through scratch in K-9. *Computers & Education*, 141, 103607. <https://doi.org/10.1016/j.compedu.2019.103607>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Mustafa Yildiz^{1,2} · Ran Shi³ · Mehmet Kara^{4,5}

Mustafa Yildiz
mustfa.yildiz@gmail.com

Ran Shi
sharon_shir@hotmail.com

- ¹ Department of Educational Sciences, Amasya University, Amasya, Turkey
- ² Department of Educational Sciences, Faculty of Education, Amasya University, Building A, Akbilek mah, Hakimiyet Caddesi No:4/3, 05100 Merkez/Amasya, Turkey
- ³ Educational Leadership Research and Technology, College of Education, Western Michigan University, 1903 West Michigan Ave., Kalamazoo, MI 49008, USA
- ⁴ Department of Computer Technology, Amasya University, Amasya, Turkey
- ⁵ Vocational School of Design, Yesilirmak Campus, Şeyhçui Mah. Kemal Nehrozoğlu Cad, 05100 Amasya, Turkey