Lee, S. W.-Y., Liang, J.-C., Hsu, C.-Y., Chien, F. P., & Tsai, M.-J. (2023). Exploring Potential Factors to Students' Computational Thinking: Interactions between Gender and ICT-resource Differences in Taiwanese Junior High Schools. *Educational Technology & Society*, *26*(3), 176-189. https://doi.org/10.30191/ETS.202307_26(3).0013

Exploring Potential Factors to Students' Computational Thinking: Interactions between Gender and ICT-resource Differences in Taiwanese Junior High Schools

Silvia Wen-Yu Lee[1], Jyh-Chong Liang[2], Chung-Yuan Hsu[3*], Francis Pingfan Chien[2] and Meng-Jung Tsai[2]

[1]Graduate Institute of Information and Computer Education and Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, Taipei, Taiwan // [2]Program of Learning Sciences and Institute for Research Excellence, National Taiwan Normal University, Taipei, Taiwan // [3]Department of Child Care, National Pingtung University of Science and Technology, Pingtung, Taiwan // swylee@ntnu.edu.tw // aljc@ntnu.edu.tw // jackohio@gmail.com // pingfan@ntnu.edu.tw // mjtsai99@ntnu.edu.tw // mjtsai99@gmail.com

[*]Corresponding author

(Submitted September 21, 2022; Revised December 26, 2022; Accepted January 16, 2023)

ABSTRACT: One of the major purposes of this study is to investigate the potential impact of gender and information and computer technology (ICT) resources on students' computational thinking (CT) competencies. To this end, the Computational Thinking Test for Junior High Students (CTT-JH) was developed and validated. Research participants included 437 junior high school students in Taiwan. The surveyed schools were categorized into *more* or *fewer* ICT resources. Then, discrimination analyses and Rasch modeling for item analyses and two-way ANOVA were conducted. Results showed that the final version of CTT-JH is of good item quality. Students in schools with more ICT resources had higher CT test mean scores regardless of gender. Nevertheless, at schools with limited resources, male students had significantly lower CT test mean scores than female students did. This study provides new insights into how gender and ICT resources can interact with and impact on students' CT competencies. It also provides a valid and reliable tool for assessing young adolescents' CT abilities.

Keywords: Computational thinking, Junior high school, Assessment, Non-programming, Domain-general CT

1. Introduction

Computational thinking (CT) can be regarded as one of the fundamental literacies of the 21st century for adapting to the future challenging society (Çoban & Korkmaz, 2021; Grover & Pea, 2018; Wing, 2006). The call for integrating CT into education has been gathering global attention during the last decade (Shute et al., 2017). Computational thinking (CT) refers to problem-solving skills (Wing, 2006) emphasizing conceptual development required to engage in formulating problems' solvable parts, abstracting key information, automating solutions through algorithmic thinking, debugging, and generalizing problem-solving processes (ISTE/ CSTA, 2011; Selby & Woollard, 2013; Shute et al., 2017). Individuals with CT skills are expected to execute the aforementioned skills to logically solve interdisciplinary and real-life problems (Araujo et al., 2019).

Various types of CT assessment instruments have been developed recently (Weintrop et al., 2021). While some instruments are to assess CT competencies based on programming and computing concepts, others are for assessing domain-specific or domain-general, non-programming problem solving competencies (Tang et al., 2020). We argued that among these different types of CT assessment, domain-general instruments that are congruent to the problem-solving perspective of CT and that can be assessed in non-computer science and even transdisciplinary learning context, require most attention from researchers and practitioners. As researchers have stressed the importance of fostering students' CT competencies at learning stages prior to college, there is a need for developing CT assessment for younger students as well (Li et al., 2020). Thus, in the current study, we developed a domain-general CT test for students at junior high school level and examined its psychometric properties. Furthermore, we are to explore the two potential factors that might impact students' CT competencies – gender and ICT-resources. The impact of gender on CT competencies remains inconclusive (e.g., Polat et al., 2021) and the relationships between ICT-resource abundance and CT competencies are by far under studied. In the following, a more comprehensive review of the different definitions of CT and the recent development of CT instruments will be introduced.

1.1. Computational thinking

A number of CT frameworks have been proposed in previous studies and these diverse frameworks imply that it is challenging to reach a consensus on CT operational definitions (Román-González et al., 2019; Shute et al., 2017; Voogt et al., 2015). Tang et al. (2020) categorized CT frameworks into two main aspects: (1) CT competencies related to programming and computing concepts, and (2) CT competencies needed for both domain-specific knowledge and general problem-solving skills. An example of the former is Brennan and Resnick's (2012) model consisting of computational concepts (sequences, conditionals, loops, etc.), practices (testing, debugging, reusing, etc.) and perspectives (viewing computation as a way of design and self-expression). An example of the latter is Shute's et al. (2017) CT model. Shute et al. (2017) synthesized 45 CT studies and proposed a competency-based CT model, not focusing on just one specific subject (e.g., coding) but approaching a problem-solving process in a systematic way. The model includes six main facets: decomposition (breaking a complicated problem into manageable parts), abstraction (identifying essential information), algorithms (logically developing solutions to a problem), debugging (finding and fixing errors), iteration (refining solutions), and generalization (transferring CT skills to other domains or situations). Similarly, Selby and Woollard (2013) reviewed CT studies and then proposed a CT model with abstraction, decomposition, algorithmic thinking, evaluations, and generalization. In sum, these frameworks can provide not only operational definitions for CT but also a foundation for CT assessments.

1.2. CT assessment

Assessments play a crucial role in determining successful integration of CT into educational contexts (Cutumisu et al., 2019; Poulakis & Politis, 2021; Tsai et al., 2022). With valid and reliable CT assessments, one can accurately evaluate students' CT development and understand the impact of the intervention (Eloy et al., 2022; Mueller et al., 2017). A majority of the assessment has been developed recently based on programming or computing concepts (i.e., the first type of instrument defined by Tang et al., 2020). For example, Román-González et al. (2018) developed Computational Thinking Test (CTt), a multiple-choice instrument for measuring learners' developmental level of CT based on fundamental programming concepts such as sequences, loops, and conditionals. Various instruments and assessment methods has been developed for measuring students' programming-based and computing-based CT. For instance, programming-based CT assessment also can be done through assessing students' programming artifacts or portfolio (Fields et al., 2021), through online puzzling games (Guenaga et al., 2021), or observing or logging students interactions (Metcalf et al., 2021). While some of the assessment utilized paper-based instruments designed for children in kindergarten in unplugged coding context (Clarke-Midura et al., 2021); others used computer automatic scoring for particular programming language, such as using Dr. Scratch for scoring Scratch-based programming artifacts (Moreno-León et al., 2015).

Nevertheless, Wing (2006) suggested that CT should not be limited to computer science or computer scientists, and further argued that CT involves computer concepts used by everyone to solve problems, manage their daily life, and interact with other people. The call for strengthening students' domain-general CT competencies has been receiving increasing attention; nevertheless, so far fewer instruments assessing domain-general CT skills and competencies are available (Tsai et al., 2021; Angeli & Giannakos, 2020; Kwon et al., 2021). Domain-general CT refers to "solving complex problems in daily life contexts" (Tsai et al., 2021, p. 2). In this sense, domain-general CT is even more important than domain-specific CT for developing future citizens' competencies for the 21st century. For instance, Tsai et al. (2022) has found that students' CT dispositions in problem-solving significantly predicted their domain-general CT competencies at elementary school level. A widely used domain-general CT assessment is the Bebras Challenge, a competition using real-life tasks to assess students' CT skills independent of previous programming experience (Dagienė & Stupuriene, 2016). The Bebras Challenge is hosted annually and internationally and more than 40 countries world-wide have participated. Example Bebras tasks can be seen at https://www.bebras.org/examples.html. Moreover, the domain-general CT instruments can be applied to various learning contexts and be utilized to examine students' CT competencies after different treatments. For instance, Chiazzese et al. (2019) measured the impacts of a robotics laboratory on the third and fourth graders' acquisition of CT competencies by using the Bebra tasks. The results showed that programming robots had a positive impact on students' acquisition of CT competencies.

Another important trend of recent research of CT assessment development is the attention to the quality of the research instruments and the scoring rubrics (Clarke-Midura et al., 2021). Researchers have raised the concerns of the lack of evaluating and reporting the validity and reliability of CT assessment in past publications (e.g., Tang et al., 2020). Using systemic methods for instrument development, such as evidence-centered design, and providing evidence of psychometric properties of instruments have been suggested by researchers when developing CT assessment (Basu et al., 2021; Clarke-Midura et al., 2021). Finally, there has been increasing attention of teaching CT for pre-college students (Weintrop et al., 2021). While recent development has shown a growth trend in CT research and CT measurement in elementary level (e.g., Basu et al., 2021; Metcalf et al., 2021; Polat et al., 2021; Tsarava et al., 2022;), the same growth has not been found at junior high school (or middle school) level. It is important to have domain-general CT instruments available at all levels for summative evaluation purposes and for monitoring students' learning progression. In the current study, we documented the process and evidence of validating a newly developed domain-general CT assessment for junior high school level (age 13-15).

1.3. Gender and digital divide

Additionally, in this study, we also aimed to explore the role of two factors in students' CT competencies–gender differences and the digital divide. Gender differences play a critical role in influencing students' CT development (Angeli & Valanides, 2020; Shute et al., 2017). Despite the fact that gender differences have been receiving growing attention recently, the findings from empirical studies seem to be inconclusive. Some studies have shown that males outperform females on CT tests at the secondary educational level (Guggemos, 2021; Tsai et al., 2022), and researchers have even reported that the higher the grade, the more intense the gender gap in CT performance (Román-González et al., 2017). Polat et al. (2021) implemented an intervention of visual programming, and found that male students tended to have better CT performance than that of girls (Polat et al., 2021). Nevertheless, in Durak and Saritepeci's (2018) study of secondary and high school students, they reported no significant relationships between gender and CT competencies.

Researchers have examined other factors, such as the type of activities, the time spent on task, or academic achievement in relation to gender differences in CT. For instance, utilizing educational robots to enhance students' CT, Angeli and Valanides (2020) found that male students benefited more from individualistic, kinesthetic, manipulative-based activities, whereas female students learned more from collaborative activities. While no statistically significant difference was found in students' CT competencies, Atmatzidou and Demetriadis (2016) found that female students required more time to reach the same CT level as males in educational robotics tasks. Furthermore, Lei et al. (2020) identified a stronger relation between CT and academic achievement among females than males in their meta-analysis research.

The digital divide is commonly defined as inequality in the use of information and communication technologies (Aydin, 2021; Light, 2009). Many researchers agree that unequal exposure to computers and advanced technology in general may impact students' interests in computer-based activities for learning and even hamper students' learning approaches and performance. Past studies have investigated the digital divide attributed to socio-economic status (SES) or the geographical location of schools. For instance, Hohlfeld et al. (2017) found that students in low-SES schools tended to use software for tutoring or practicing, while those from high-SES schools were inclined to use software more for researching, communication, and developing projects to demonstrate what they had learned. Moreover, Zhang (2014) utilized Google Trends and Web analytics to investigate middle and elementary school students' usage of the PhET website, one of the most well-known online science simulation resources. The results showed that students in high SES families were more interested in using PhET for learning sciences than their low-SES counterparts. In terms of geographical location, Kale et al. (2018) found that school rurality may influence teachers' own CT competence and their teaching of CT in classrooms. In other words, rural primary school teachers tended to have limited CT skills and felt that they were not ready to integrate CT into their teaching.

In the current study we examined the digital divide by using the schools' information and computer technology (ICT) resources as an indicator rather than school location. It is our observation that in Taiwan, school location does not necessarily contribute to the abundance or lack of ICT resources. In other words, rural schools or county-funded schools may have equal or more ICT resources than urban schools if they are enlisted as one of the ICT-schools or if the school is ambitious in getting more funding for ICT.

1.4. Purpose

Although past studies have shown the impact of digital divide on ICT competence or ICT attitudes, few studies, on the one hand, have examined its impact on students' CT competence or CT perceptions. As computational thinking is a 21st century skill in the current technological world, gaining more insights into how the digital divide influences students' CT development has become essential (Czerkawski & Lyman, 2015). On the other hand, it remains inconclusive under which conditions gender differences existed in CT performance. While gender and the digital divide are important issues in ICT literacy (e.g., Kim et al., 2021), one of the major purposes of this study is to investigate the potential impact of the gender and the ICT-resources, and their interactions on students' CT competencies.

To this end, it is important to have a valid, reliable, age-appropriate, and domain-general CT assessment instrument for researchers and teacher. A CT assessment tool, the Computational Thinking Test for Junior-High School Students (CTT-JH) was developed and validated in this study. In this study, the test items were adapted from the items from the Bebras Challenge and we revised the language, context, and presentation to make it suitable for students in Taiwan. The Bebras Challenge has been adopted internationally and has reached success in promoting computational thinking worldwide, however, only a few studies have examined its psychometric properties for research purposes. While some studies have used content analysis or success rate for analyzing item difficulty (Izu et al., 2017; van der Vegt, 2018), we suggested using Rasch modeling based on Item Response Theory (IRT) to provide more rigorous evidence regarding item quality. In sum, we posed the following research questions: (1) what are the validity and reliability of CTT-JH? (2) What are the effects of gender and ICT-resources on students' CT competencies?

2. Methods

2.1. Participants

Participants of the current study were 437 junior high school students in Taiwan, including 234 males (about 53.5%) and 203 females. Among the participants 105 were seventh grade students (about 24.0%), 162 eighth graders (about 37.1%), and 170 ninth graders (about 38.9%). The students were recruited from 16 intact classes in six junior high schools (two from a city, one from a county, two from a rural area, and one from a remote area) in the north and the center of Taiwan. To meet the ethical requirements, the participants were informed that their involvement in the study was voluntary and that their personal information would be treated confidentially. They were informed that they may withdraw from the study at any time. The students who agreed to participate then complete the CTT-JH items within an hour. The response rate was about 97.1%. School ICT resource information was collected from the school ICT administrator. All participants were assumed to have problem solving experience in their daily lives as well as in academic learning domains such as mathematics and science. They also had experience of using ICT for learning before participation.

The six schools were divided to more or fewer ICT resource groups based on the following three criteria : (1) the ratio of full-time ICT teacher to class, if the ratio for a school was greater than 0.1 then the school was coded as 1, otherwise was coded as 0 ; (2) the funding of Maker Education and Technology Center from the government, the schools with the findings were coded as 1, otherwise was coded as 0; and (3) the implementation of project-based ICT-integrated curriculum, the school that conducted the curriculum was coded as 1, otherwise was coded as 0. These criteria indicated the likelihood for students to be taught by full-time ICT teachers, the school involvement in maker education, and their implementation of ICT-integrated curriculum. Data were obtained from the Ministry of Education of Taiwan during 2019-2020. After the coding, we found two schools had ICT-teacher to class ratio more than 0.1; three schools received funding during 2019-2020 for maker centers; and two schools had projects for designing ICT-integrated curriculum. By summing up the three indices, each school obtained a total score that indicated the ICT resource of the school. If a school received a total score of 0, then the school was categorized into the fewer-resource group; otherwise it was categorized into the more-resource group. For detailed information of each school, please see Appendix. Finally, four schools were labeled as more-resource schools and two schools were labeled as fewer-resource schools.

2.2. Research instrument

The Computational Thinking Test for Junior High students (CTT-JH) was a test developed to measure junior high school students' CT performance in this current study. Revised from the Bebras Challenge tasks, a pool of 15 items built up the initial version of the CTT-JH for assessing abstraction, decomposition, algorithmic thinking, evaluation, and generalization. We also referred to the CT framework in which these five dimensions respectively refer to the ability to abstract essential information, to break down complicated problems into manageable parts, to think procedurally as a sequence of steps to reach a solution, to decide the most appropriate solution to the problem, as well as to adapt and transfer solutions to other problems. Each item was designed to assess one or more CT dimensions simultaneously, and collectively CTT-JH is a multi-dimensional research instrument. All the items were redesigned or modified as solving problems in a daily-life farm-based context in

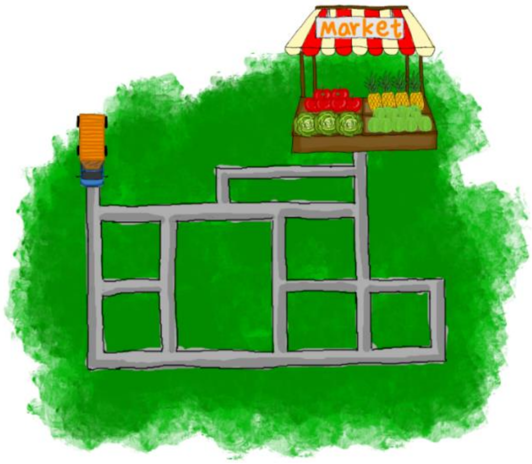
Taiwan. Two sample items are illustrated in Figure 1 and the complete test is available online at https://bit.ly/2022-CTT-JH.

Figure 1. Sample items of CTT-JH

Q2. Selling goose eggs (easy)

A truck from Jack’s Happy Farm is on the way to the market to sell goose eggs. The truck can drive in only 3 ways:

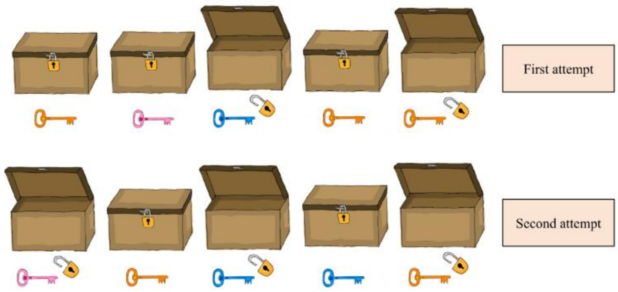
1. Left: Turn 90 degrees left
2. Right: Turn 90 degrees right
3. Forward: Go forward until you cannot go forward anymore



Question: Write a set of instructions (a program) that will get the truck to the market. You can do this by writing down the code numbers.

Q4. Hungry goose (difficult)

On Jack’s Happy Farm, a hungry goose is trying to unlock five food boxes. Jack gives the goose 3 keys of different colors and says that these keys can open all the boxes. The results of the goose’s first and second attempts are shown below.



Question: Which one is the correct order of the keys to open all the boxes?

1. Blue, Pink, Blue, Orange, Orange
2. Pink, Blue, Blue, Blue, Orange
3. Pink, Blue, Blue, Pink, Orange
4. Pink, Pink, Blue, Pink, Orange

2.3. Data analysis

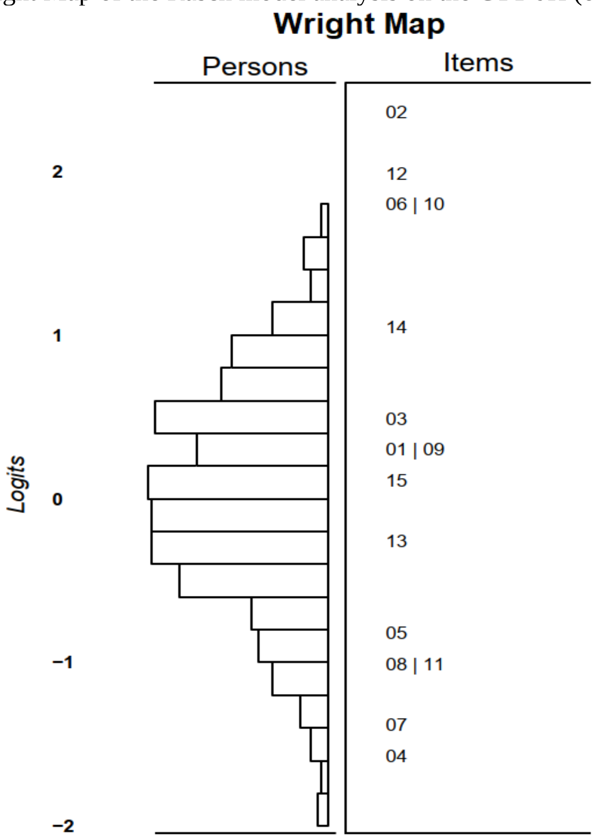
To ensure the content validity of the CTT-JH, the items were reviewed by the research team who had expertise in computer education, educational technology, and science education. Through meetings, consensus about the CT dimensions assessed by each CTT-JH item was developed among the experts. A list of items and its corresponding CT dimensions will be presented in the result section. To understand whether the CTT-JH test items are fitted and reliable measurement for junior high school students, we applied the Rasch model, a one-parameter logistic Item Response Theory (IRT) model for dichotomous items (Andrich & Marais, 2019; Mayer et al., 2014; Rasch, 1960) for data analysis. Test Analysis Modules (TAM) and the Wright Map packages in R software were used to estimate item difficulties and students’ abilities on the same logit scale (Robitzsch et al., 2020; Irribarra & Freund, 2014). Finally, in order to examine whether there was any significant differences in the participants' CTT-JH test scores due to gender and teaching resources, a 2 x 2 ANOVA (gender x resource) was conducted.

3. Results

3.1. Psychometric properties of the CTT-JH items

Figure 2 illustrates the Wright Map of the Rasch model analysis of the participants' CTT-JH scores. It shows the distributions of the student's abilities (on the left side) and the distributions of the item difficulties (on the right side). The original 15 items were ordered from the most difficult (at the top, i.e., item 2) to the least difficult (at the bottom, i.e., item 4). The histograms of student's abilities show that each student solved the item with a probability of 50% and are plotted from most able (at the top) to least able (at the bottom).

Figure 2. The Wright Map of the Rasch model analysis on the CTT-JH (original 15 items)



After we fitted the Rasch model for the original 15 items, we examined the reliability of the whole test. The Weighted Likelihood Estimate (WLE) person-separation reliability was 0.53 and an Expected A Posteriori estimate based on Plausible Values (EAP/PV) reliability was 0.56, which was slightly lower than the acceptable value of 0.6. This suggested that some of the items in the original test might need to be reconsidered for inclusion in the test.

Table 1 summarizes the item properties of the IRT Rasch model and of the classical discriminant analysis. The items are listed from Q2 (i.e., item 2) to Q4 (i.e., item 4) according to their item difficulties ranging from 2.42 (most difficult) to -1.64 (least difficult) as well as their correct response rates ranging from 10% (lowest) to 81.33% (highest). The average person's proficiency was 0.00004 logits (*SD* = 0.93). The fit for single items (weighted mean squares, MNSQ) ranged from 0.89 to 1.12 (Mean = 1.00, *SD* = 0.06), thus indicating a good fit to the Rasch model at the item level. Finally, we applied point biserial correlations for the correct answers to obtain the classical discrimination values that ranged from 0.03 to 0.60.

In order to improve the reliability of the original version of the CTT-JH, each item was carefully examined based on the data reported in Table 1. First, Q10 was deleted due to its extremely low discrimination (0.03). Then, each of the remaining items was checked to ascertain whether the overall reliability would be increased when it was deleted. Finally, the best acceptable reliability of the overall test (EAP/PV reliability = 0.61) was obtained when Q12 and Q3 were deleted. Therefore, after deleting the three items Q10, Q12, and Q3, the final version of CTT-JH was formed with 12 items, as shown in Figure 3 and Table 2.

Table 1. Item properties of the Rasch model analysis and classical discrimination on the Junior High school computational thinking test (original 15 candidate items)

Item	Percent correct (%)	Difficulty	Discrimination	Infit MNSQ
Q2	10.00	2.42	0.35	0.97
Q12	14.44	1.98	0.23	1.03
Q10	16.00	1.81	0.03	1.12
Q6	16.89	1.76	0.39	0.97
Q14	29.11	1.00	0.28	1.05
Q3	39.78	0.47	0.30	1.06
Q9	43.56	0.29	0.40	1.00
Q1	44.67	0.24	0.37	1.02
Q15	47.78	0.10	0.39	1.02
Q13	55.11	-0.23	0.42	1.00
Q5	66.67	-0.78	0.60	0.89
Q11	72.44	-1.09	0.54	0.92
Q8	73.11	-1.12	0.52	0.92
Q7	79.33	-1.51	0.33	1.00
Q4	81.33	-1.64	0.43	0.96

Note. EAP/PV reliability = 0.56, WLE reliability = 0.53.

Figure 3. Wright Map of Rasch analysis on CTT-JH (12 items)

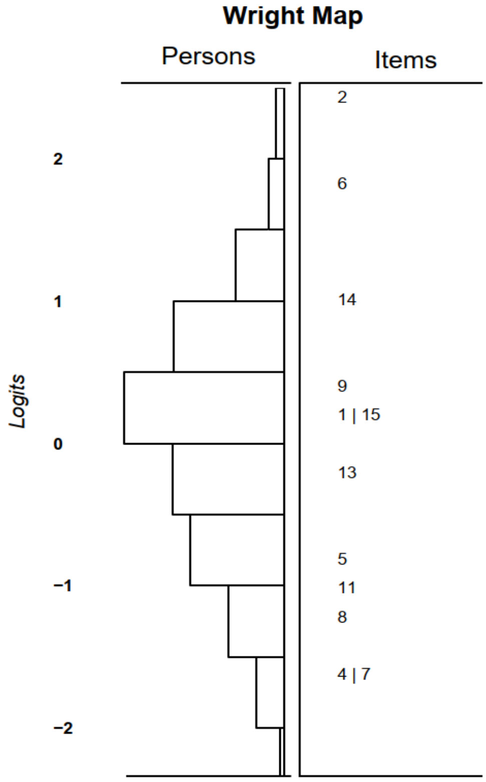


Figure 3 displays the distribution of students' abilities (on the left side) and item difficulties (on the right side) on the same logit scale. Items are ordered from the most difficult (at the top) to the least difficult (at the bottom). The histograms of students' abilities show that each student solved the item with a probability of 50% and was plotted from most able (at the top) to least able (at the bottom).

After we fitted the Rasch model, the results showed that the Expected A Posteriori estimate based on Plausible Values (EAP/PV) reliability was 0.61, Weighted Likelihood Estimate (WLE) person-separation reliability was 0.57, and Cronbach's alpha reliability was .6. Item difficulties ranged from 2.47 to -1.68. The average person's proficiency is 0.00018 logits (*SD* = 1.12). The fit for single items (weighted mean squares, MNSQ) ranged from 0.88 to 1.08 (Mean = 1.00, *SD* = 0.06), thus indicating a good fit to the Rasch model at the item level. In addition, we applied point biserial correlations for the correct answers to obtain the classical discrimination values that ranged from 0.28 to 0.60. The final list of items and its corresponding CT dimensions are shown in Table 3. The assessment items and the answering keys are available online at https://bit.ly/2022-CTT-JH.

Table 2. Item properties of Rasch analysis and classical discrimination on the final version of the CTT-JH (final 12 items)

Item	Percent correct (%)	Difficulty	Discrimination	Infit MNSQ
Q2	10.00	2.47	0.35	0.99
Q6	16.89	1.82	0.39	1.00
Q14	29.11	1.03	0.28	1.08
Q9	43.56	0.30	0.40	1.05
Q1	44.67	0.25	0.37	1.06
Q15	47.78	0.10	0.39	1.04
Q13	55.11	-0.24	0.42	1.01
Q5	66.67	-0.80	0.60	0.88
Q11	72.44	-1.11	0.54	0.91
Q8	73.11	-1.15	0.52	0.92
Q7	79.33	-1.54	0.33	1.03
Q4	81.33	-1.68	0.43	0.98

Note. EAP/PV reliability = 0.61, WLE reliability = 0.57, Cronbach's alpha = 0.6.

Table 3. Items of the CTT-JH responding to Selby and Woollard's (2013) CT framework

Item	Decomposition	Abstraction	Algorithm	Evaluation	Generalization
Q1. A toy goose is going out of farm			V		
Q2. The way to sell goose eggs			V		
Q4. Hungry goose open the boxes		V		V	
Q5. Transforming goose		V		V	
Q6. Let's shake hands after the match	V				
Q7. Best place for a bus stop				V	
Q8. Navigation app		V		V	V
Q9. Jack's code, QJ-Code	V				V
Q11. Swap the order and tell the secret			V		
Q13. The vine's weekly growth		V			V
Q14. Jack's self-driving car			V		V
Q15. Jack's henhouse management				V	V
Total items per dimension	2	5	4	4	4

3.2. The potential impact of gender and school resources on students' CT scores

Two-Way ANOVA was conducted to evaluate the effects of gender and different school ICT resources on the CT test mean scores. Homogeneity of variance of the four groups was verified according to Howell's study (2013, p. 234), which indicated that the results of variance analysis were more likely to be valid when the ratio of largest variance to smallest variance was four or below among the groups. In the current study, the ratio was 2.05 and revealed that the homogeneity assumption was not violated. Table 4 summarizes the two-way ANOVA results. No main effect for gender was observed ($F = 1.71$, $p > .05$, Partial eta squared < 0.01). However, school ICT resources reached significance on the CT test mean scores with medium to large effect size ($F = 41.76$, $p < .01$, Partial eta squared = 0.09). Most importantly, significant interaction with small to medium effect size occurred between gender and school ICT resources on the CT test mean scores ($F = 5.86$, $p < .05$, Partial eta squared = 0.01).

The regression line of students' school with more or fewer resources on CT mean scores was plotted for the different gender groups to better explain the interaction effects of Gender*School resources on CT test mean scores, as shown in Figure 4. To follow up on the significant interaction and to examine the differences between male and female students at different school resource levels, descriptive statistics and two independent *t* tests were used. Table 5 displays the results.

Table 4. Two-way ANOVA of the CT test scores

Source	df	MS	*F*	Partial eta squared
Gender	1	0.06	1.71	< 0.01
Schools with more or fewer ICT resources	1	1.35	41.76**	0.09
Gender * School ICT resources	1	0.19	5.86*	0.01

Note. df = degree of freedom; MS = Mean squares; $^{**}p < .01$; $^{*}p < .05$.

Firstly, the plots revealed that students' school level had a positive association with their CT test mean scores; that is, students in better resourced schools had higher CT test mean scores regardless of gender. Second, at schools with limited resources, male students had significantly lower CT test mean scores than female students did (male students' M = 0.41, SD = 0.20; female students' M = 0.48, SD = 0.14; t value = -2.37, $p < .05$). However, there was no significant difference in CT test mean scores between male and female students in schools with more resources (male students' mean scores = 0.57, standard deviation = 0.19; female students' mean scores = 0.55, standard deviation = 0.17; t value = 0.97, $p > .05$). Thirdly, the different slopes revealed that school ICT resources had a greater effect on students' CT test mean scores for male students than for female students. To summarize, it can be stated that school ICT resources, as well as the relationship of students' school ICT resources and gender, may be critical variables for CT test mean scores.

Figure 4. Interaction effect of gender and schools with more or fewer resources

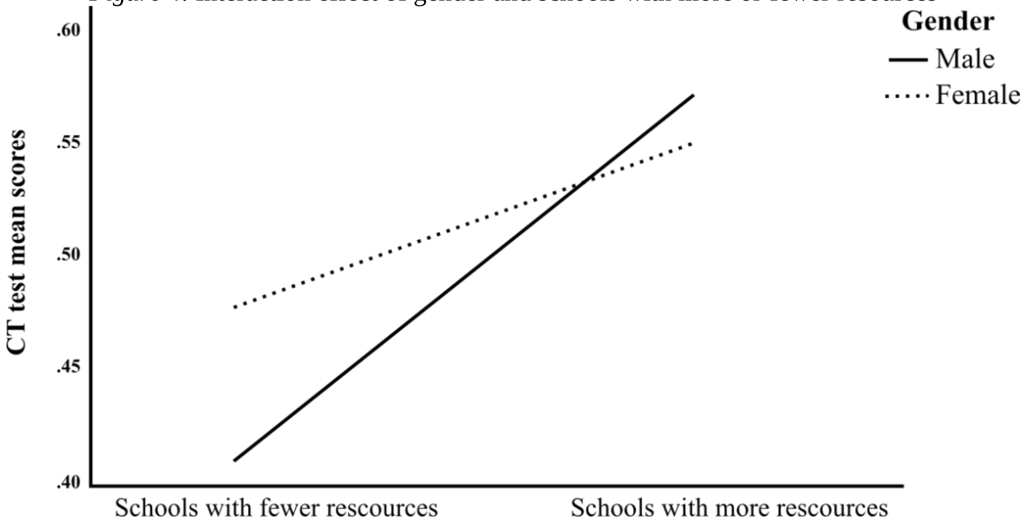


Table 5. Descriptive statistics for the independent variables

School levels	Gender	Mean	*SD*	*N*	*t*-test
Fewer resources	Male	0.41	0.20	77	-2.37*
	Female	0.48	0.14	65	
More resources	Male	0.57	0.19	157	0.97
	Female	0.55	0.17	138	

Note. $^{*}p < .05$.

4. Conclusions and discussion

The present study aimed to develop and validate a computational thinking test for junior high students. The results of IRT analysis showed that the revised version of the CTT-JH test is a reliable instrument for measuring junior high school students' computational thinking. While the Bebras Challenge tasks have been used worldwide, researchers have pointed out problems with the quality of the items (Hubwieser & Muhling, 2015). Nevertheless, only a few studies have provided robust evidence of the psychometric properties of this type of domain-general CT assessment instrument. Results of the current study show that through iteratively using the IRT Rasch model and the classical discriminant analysis, we were able to identify and remove unsuitable items. Our revised Bebras Challenge items, with attention to the wording, the representation, and the new context, represent a joint construct of domain-general CT. The final version including 12 items is suitable for measuring the CT competencies of students at junior-high school level.

The CTT-JH research instrument has potential applications and research implications in future studies. First, it can be used in both pretests and posttests as almost no prior instruction in particular discipline or logic training required. This paper-based instrument can also be adapted to assess students' learning gains in computer-free CT learning activities such unplugged computing robots. Second, because of the domain general nature this instrument, it is possible to assess students' CT across different disciplines and even in a transdisciplinary learning context. Some researchers have conceptualized itself CT as a transdisciplinary concept and the needs of integrative thinking skills (Li et al., 2020). While no commonly accepted definitions of integrative thinking are available and instruments for assessing integrative thinking are scarce (National Research Council, 2014), we argued that domain-general CT assessment can be used for measuring integrative thinking when it is used in transdisciplinary context such as STEM education.

The results also show that school ICT resources, as well as the interaction between school ICT resources and gender, may be critical variables for students' domain-general CT competencies. On the one hand, students in schools with more ICT resources had higher CT test mean scores regardless of gender. This finding not only supports prior research which found no gender differences in CT test results (Durak & Saritepeci, 2018), but more importantly, it implies that school ICT resources play some role in students' development of CT competencies even for the non-programming, domain-general CT. One possible explanation is that digital learning nowadays has been applied to different aspects of learning. When students study in an advanced ICT environment, they might have more access to ICT use for various cognitive tasks such as analyzing, creating, exchanging, and using data and information in different subject areas (Herselman & Britton, 2002). In another study, Sirakaya (2020) found that students' CT skills were associated with their internet experience, mobile device experience, and mobile internet experience. These findings support the association between CT competencies and ICT usage. Directly or indirectly, access to and use of ICT resources may have helped to develop students' domain-general CT and to close the digital divide (Rallet & Rochelandet, 2007). An important implication of this finding is that school ICT resources do not only impact students' computer literacy or programming learning, but may also influence students' domain-general CT competencies as one key competency in the 21st century. This is an important area that should be considered in future educational policy for school ICT funding.

On the other hand, at schools with limited resources, male students had significantly lower CT test mean scores than female students did. In other words, male more than female students' CT competencies are affected by the lack of ICT resources in schools. Attention to the interactions between gender and ICT provides another angle for possible explanations of why empirical evidence of gender differences was inconclusive. We hypothesized that there might be different models of how students develop CT in ICT-deprived versus ICT-advanced environments. In ICT-deprived learning environments, students' domain-general CT competencies might have strong relationships to academic achievement. Previous research reported that other academic skills such as mathematical thinking and reading and verbal skills (Zhang & Nouri, 2019; Roman-Gonzalez et al., 2018) were found to be related to students' CT competencies. Furthermore, in a previous meta-analysis study (Lei et al., 2020), researchers concluded that students' CT is correlated to school achievement; furthermore, the correlations are stronger among female than male students. In other words, in ICT-deprived schools, female students' better CT competencies than male students may be related to female students' overall school achievement.

Another possible explanation regards the gender differences of ICT usage outside of schools. Kim et al. (2021) surveyed 23,000 elementary and middle school students in Korea and found that female students had higher ICT literacy levels than male students. They attributed the ICT literacy difference to the different ICT usage habits and attitudes of males and females. For instance, researchers have found gender differences in Internet using purposes and intensity at high school level (Tsai & Tsai, 2010); female students tended to use the Internet for communication purposes while male students tended to use the Internet for exploration purposes. Moreover, female students are more likely to use ICT after school for learning or doing homework than male students (Ahn & Chae, 2016), and female students used ICT to gain more experience of problem solving through social networks while male students used ICT for entertainment and games (Sung & Choi, 2016). The aforementioned studies provide possible explanations as to why in the current study we found that female students outperformed male students. When schools have fewer ICT resources, students' habits of ICT usage outside school can become even more influential. How to help male students to gain ICT competencies in ICT-deprived learning environments and how to gain understanding of what causes the gender differences are important questions to be studied in the future.

Finally, we identified some limitations in this study. In the current study, we explored the impact of gender and school resource interactions but did not have data to identify the epistemic resources of students' CT competencies. Careful and in-depth inquiries into the gender differences in conjunction with school ICT and comprehensive data collection, such as including data of ICT usage in-class and outside of school are suggested for future research. Moreover, we were able to categorize the students' schools into fewer or more ICT resources by reviewing ICT-related information from the schools. Future studies can further develop a system to quantify the school-level ICT information or quantify student-level ICT usage and include the data in a more complex statistical model by using statistics such as hierarchical linear modeling (HLM). Perhaps it is not possible to fully understand the relationships among domain-general ICT, gender, and ICT resources without expanding the understanding to students' other academic competencies or students' ICT usage in daily life. Further investigations of this area are required in future research.

Acknowledgment

This study was supported by the Ministry of Science and Technology (MOST) of Taiwan under the following project numbers: MOST 106-2511-S-003-065-MY3, 109-2511-H-003-019-MY3, 109-2511-H-003-052-MY3, 109-2628-H-020-001-MY3 and 111-2410-H-003-130-MY3. It was also supported by the "Institute for Research Excellence in Learning Sciences" of NTNU sponsored by the Ministry of Education (MOE) in Taiwan.

References

Ahn, S. & Chae, K. (2016). Correlation analysis on ICT literacy level and difference of habit to use ICT. *Journal of the Korean Association of Information Education, 20*(3), 303-312. https://doi.org/10.14352/ jkaie.2016.20.3.303

Andrich, D., & Marais, I. (2019). *A Course in Rasch measurement theory*. Springer.

Angeli, C., & Giannakos, M. (2020). Computational thinking education: Issues and challenges. *Computers in Human Behavior, 105,* 106185. https://doi.org/10.1016/j.chb.2019.106185

Angeli, C., & Valanides, N. (2020). Developing young children's computational thinking with educational robotics: An Interaction effect between gender and scaffolding strategy. *Computers in Human Behavior, 105,* 105954. https://doi.org/10.1016/j.chb.2019.03.018

Araujo, A. L. S. O., Andrade, W. L., Guerrero, D. D. S., & Melo, M. R. A. (2019). How many abilities can we measure in computational thinking? A Study on Bebras challenge. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (pp. 545–555). https://doi.org/10.1145/3287324.3287405

Atmatzidou, S., & Demetriadis, S. (2016). Advancing students' computational thinking skills through educational robotics: A Study on age and gender relevant differences. *Robotics and Autonomous Systems, 75,* 661-670. https://doi.org/10.1016/j.robot.2015.10.008

Aydin, M. (2021). Does the digital divide matter? Factors and conditions that promote ICT literacy. *Telematics and Informatics, 58,* 101536. https://doi.org/10.1016/j.tele.2020.101536

Basu, S., Rutstein, D. W., Xu, Y., Wang, H., & Shear, L. (2021). A Principled approach to designing computational thinking concepts and practices assessments for upper elementary grades. *Computer Science Education, 31*(2), 169-198. https://doi.org/10.1080/08993408.2020.1866939

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. 25).

Chiazzese, G., Arrigo, M., Chifari, A., Lonati, V., & Tosto, C. (2019). Educational robotics in primary school: Measuring the development of computational thinking skills with the Bebras tasks. *Informatics, 6*(4), 43. https://doi.org/10.3390/informatics6040043

Clarke-Midura, J., Silvis, D., Shumway, J. F., Lee, V. R., & Kozlowski, J. S. (2021). Developing a kindergarten computational thinking assessment using evidence-centered design: The Case of algorithmic thinking. *Computer Science Education, 31*(2), 117-140. https://doi.org/10.1080/08993408.2021.1877988

Çoban, E., & Korkmaz, Ö. (2021). An Alternative approach for measuring computational thinking: Performance-based platform. *Thinking Skills and Creativity, 42,* 100929. https://doi.org/10.1016/j.tsc.2021.100929

Cutumisu, M., Adams, C., & Lu, C. (2019). A Scoping review of empirical research on recent computational thinking assessments. *Journal of Science Education and Technology, 28*(6), 651-676. https://doi.org/10.1007/s10956-019-09799-3

Czerkawski, B. C., & Lyman, E. W. (2015). Exploring issues about computational thinking in higher education. *TechTrends, 59*(2), 57-65. https://doi.org/10.1007/s11528-015-0840-3

Dagienė, V., & Stupuriene, G. (2016). Bebras - A Sustainable community building model for the concept based learning of informatics and computational thinking. *Informatics in Education, 15*(1), 25-44. https://doi.org/10.15388/infedu.2016.02

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

Eloy, A., Achutti, C. F., Fernandez, C., & Lopes, R. D. (2022). A Data-driven approach to assess computational thinking concepts based on learners' artifacts. *Informatics in Education, 21*(1), 33-54. https://doi.org/10.15388/infedu.2022.02

Fields, D., Lui, D., Kafai, Y., Jayathirtha, G., Walker, J. & Shaw, M. (2021) Communicating about computational thinking: Understanding affordances of portfolios for assessing high school students' computational thinking and participation practices. *Computer Science Education, 31*(2), 224-258. https://doi.org/10.1080/08993408.2020.1866933

Grover, S., & Pea, R. (2018). Computational thinking: A Competency whose time has come. In *Computer science education: Perspectives on teaching and learning in school* (pp. 19-38). Bloomsbury Academic.

Guenaga, M., Eguíluz, A., Garaizar, P. & Gibaja, J. (2021) How do students develop computational thinking? Assessing early programmers in a maze-based online game. *Computer Science Education, 31*(2), 259-289. https://doi.org/10.1080/08993408.2021.1903248

Guggemos, J. (2021). On the predictors of computational thinking and its growth at the high-school level. *Computers & Education, 161,* 104060. https://doi.org/10.1016/j.compedu.2020.104060

Herselman, M., & Britton, K. G. (2002). Analysing the role of ICT in bridging the digital divide amongst learners. *South African Journal of Education*, *22*(4), 270-274.

Hohlfeld, T. N., Ritzhaupt, A. D., Dawson, K., & Wilson, M. L. (2017). An Examination of seven years of technology integration in Florida schools: Through the lens of the levels of digital divide in schools. *Computers & Education, 113,* 135-161. https://doi.org/10.1016/j.compedu.2017.05.017

Howell, D. C. (2013). *Statistical methods for psychology*. Wadsworth Cengage Learning.

Hubwieser, P., & Mühling, A. (2015). Investigating the psychometric structure of Bebras contest: Towards measuring computational thinking skills. In P*roceedings of 2015 International Conference on Learning and Teaching in Computing and Engineering*. https://doi.org/10.1109/LaTiCE.2015.19

Irribarra, D. T., & Freund, R. (2014). *Wright Map: IRT item-person map with ConQuest integration.* http://github.com/david-ti/wrightmap.

International Society for Technology in Education (ISTE) & Computer Science Teachers Association (CSTA). (2011). *Operational definition of computational thinking for K-12 education*. https://cdn.iste.org/www-root/Computational_Thinking_Operational_Definition_ISTE.pdf

Izu, C., Mirolo, C., Settle, A., Mannila, L., & Stupurienė, G. (2017). Exploring Bebras tasks content and performance: A Multinational study. *Informatics in Education, 16*(1), 39-59. http://dx.doi.org/10.15388/infedu.2017.03

Kale, U., Akcaoglu, M., Cullen, T., & Goh, D. (2018). Contextual factors influencing access to teaching computational thinking. *Computers in the Schools, 35*(2), 69-87. https://doi.org/10.1080/07380569.2018.1462630

Kim, H. S., Kim, S., Na, W., & Lee, W. J. (2021). Extending computational thinking into information and communication technology literacy measurement: Gender and grade issues. *ACM Transactions on Computing Education (TOCE)*, *21*(1), 1-25. https://doi.org/10.1145/3427596

Kwon, K., Cheon, J., & Moon, H. (2021). Levels of problem-solving competency identified through Bebras Computing Challenge. *Education and Information Technologies, 26*(5), 5477-5498. https://doi.org/10.1007/s10639-021-10553-9

Lei, H., Chiu, M. M., Li, F., Wang, X., & Geng, Y. J. (2020). Computational thinking and academic achievement: A Meta-analysis among students. *Children and Youth Services Review, 118*, 105439. https://doi.org/10.1016/j.childyouth.2020.105439

Li, Y., Schoenfeld, A. H., diSessa, A. A., Graesser, A. C., Benson, L. C., English, L. D., & Duschl, R. A. (2020). On computational thinking and STEM education. *Journal for STEM Education Research, 3*(2), 147-166. https://doi.org/10.1007/s41979-020-00044-w

Light, J. (2009). Rethinking the digital divide. *Harvard Educational Review, 71*(4), 709-734. https://doi.org/10.17763/haer.71.4.342x36742j2w4q82

Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction, 29*, 43–55.

Metcalf, S. J., Reilly, J. M., Jeon, S., Wang, A., Pyers, A., Brennan, K., & Dede, C. (2021). Assessing computational thinking through the lenses of functionality and computational fluency. *Computer Science Education, 31*(2), 199-223. https://doi.org/10.1080/08993408.2020.1866932

Moreno-León, J., Robles, G., & Román-González, M. (2015). Dr. Scratch: Automatic analysis of scratch projects to assess and foster computational thinking. RED. *Revista de Educación a Distancia,* (46), 1-23.

Mueller, J., Beckett, D., Hennessey, E., & Shodiev, H. (2017). Assessing computational thinking across the curriculum. In P. J. Rich & C. B. Hodges (Eds.), *Emerging research, practice, and policy on computational thinking, educational* *communications and technology: Issues and innovations* (pp. 251–267). Springer. https://doi.org/10.1007/978-3-319-52691-1_16

National Research Council. (2014). *STEM integration in K-12 education: Status, prospects, and an agenda for research.* National Academies Press. https://doi.org/10.17226/18612

Polat, E., Hopcan, S., Kucuk, S., & Sisman, B. A. (2021). Comprehensive assessment of secondary school students' computational thinking skills. *British Journal of Educational Technology*, 52, 1965-1980. https://doi.org/10.1111/bjet.13104

Poulakis, E., & Politis, P. (2021). Computational thinking assessment: Literature review. In T. Tsiatsos, S. Demetriadis, A. Mikropoulos, & V. Dagdilelis (Eds.), *Research on E-Learning and ICT in Education: Technological, Pedagogical and Instructional Perspectives* (pp. 111-128). Springer International Publishing. https://doi.org/10.1007/978-3-030-64363-8_7

Rallet, A., & Rochelandet, F. (2007). ICTs and inequalities: The Digital divide. In E. Brousseau, & N. Curien (Eds.), *Internet and digital economics: principles, methods and applications* (pp. 693–717). Cambridge University Press. https://doi.org/10.1017/cbo9780511493201.025

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Institute of Educational Research.

Robitzsch, A., Kiefer, T., & Wu, M. (2020). *TAM: Test analysis modules. R package version 2.6-2*. http://CRAN.R-project.org/package=TAM

Román-González, M., Moreno-León, J., & Robles, G. (2019). Combining assessment tools for a comprehensive evaluation of computational thinking interventions. In S.-C. Kong & H. Abelson (Eds.), *Computational Thinking Education* (pp. 79-98). Springer Singapore.

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

Román-González, M., Pérez-González, J.-C., Moreno-León, J., & Robles, G. (2018). Extending the nomological network of computational thinking with non-cognitive factors. *Computers in Human Behavior, 80,* 441-459. https://doi.org/10.1016/j.chb.2017.09.030

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

Sirakaya, D. A. (2020). Investigation of computational thinking in the context of ICT and mobile technologies. *International Journal of Computer Science Education in Schools, 3*(4), 50-59. https://doi.org/10.21585/ijcses.v3i4.73

Sung, E. & Choi, H. (2016). Exploring the types of classes and characteristics on ICT literacy of middle school students with latent class analysis. *Journal of Educational Technology, 32*(4), 987-1013. http://doi. org/ 10.17232/KSET.32.4.987

Tang, X., Yin, Y., Lin, Q., Hadad, R., & Zhai, X. (2020). Assessing computational thinking: A Systematic review of empirical studies. *Computers & Education, 148,* 103798. https://doi.org/10.1016/j.compedu.2019.103798

Tsai, M.-J., & Tsai, C.-C. (2010). Junior high school students' Internet usage and self-efficacy: A Re-examination of the gender gap. *Computers & Education, 54*(4), 1182-1192. https://doi.org/10.1016/j.compedu.2009.11.004

Tsai, M.-J., Chien, F. P., Lee, S. W.-Y., Hsu, C.-Y., & Liang, J.-C. (2022). Development and validation of the computational thinking test for elementary school students (CTT-ES): Correlate CT competency with CT disposition. *Journal of Educational Computing Research, 60*(5), 1110-1129. https://doi.org/10.1177/07356331211051043

Tsai, M.-J., Liang, J.-C., & Hsu, C.-Y. (2021). The Computational thinking scale for computer literacy education. *Journal of Educational Computing Research, 59*(4), 579-602. https://doi.org/10.1177/0735633120972356

Tsarava, K., Moeller, K., Román-González, M., Golle, J., Leifheit, L., Butz, M. V., & Ninaus, M. (2022). A Cognitive definition of computational thinking in primary education. *Computers & Education, 179,* 104425. https://doi.org/10.1016/j.compedu.2021.104425

van der Vegt, W. (2018). How hard will this task be? Developments in analyzing and predicting question difficulty in the Bebras Challenge. *Olympiads in Informatics, 12,* 119-132. http://dx.doi.org/10.15388/ioi.2018.10

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

Weintrop, D., Rutstein, D. W., Bienkowski, M., & Mcgee, S. (2021). Assessing computational thinking: An Overview of the field. *Computer Science Education, 31*(2), 113-116. https://doi.org/10.1080/08993408.2021.1918380

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33-35.

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

Zhang, M. (2014). Who are interested in online science simulations? Tracking a trend of digital divide in Internet use. *Computers & Education*, *76,* 205-214. https://doi.org/10.1016/j.compedu.2014.04.001

Appendix

ICT resource information for each school where data were collected

School[a]	Location	(a) Number of full-time ICT teacher in school	(b) Number of classes in school	(c) = (a) / (b) Full-time ICT teacher-class ratio	(1) ICT teacher-class ratio (if c > 0.1 coded 1, otherwise coded 0)	(2) Funding of Maker Education and Technology Center (Yes = 1, Not Available = 0)	(3) ICT-integrated curriculum project (Yes=1, Not Available=0)	Total Score = (1)+(2)+(3)	ICT resources in school (More: Total Score > 0, Fewer: Total Score = 0)
School A	City	3	61	0.049	0	1	0	1	More resource
School B	City	2	33	0.061	0	0	0	0	Fewer resource
School C	County	0	24	0.000	0	0	0	0	Fewer resource
School D	Rural	3	27	0.111	1	0	1	2	More resource
School E	Rural	6	43	0.140	1	1	0	2	More resource
School F	Remote	0	28	0.000	0	1	1	2	More resource

Note. [a]Data were obtained from the Ministry of Education of Taiwan during 2019