

Students' Computational Thinking Skills and Intentions to Continue Their Studies in the IT Field

Riin SAADJÄRV* [0000-0003-1275-1090], Piret LUIK [0000-0002-9049-4192], Tauno PALTS [0000-0001-7494-837X]

Institute of Computer Science, University of Tartu, Estonia

e-mail: riin.saadjarv@ut.ee, piret.luik@ut.ee, tauno.palts@ut.ee

Abstract. The development of students' computational thinking (CT) skills is a crucial aspect in today's world. We used 12 Bebras-based items mapped to combinations of four CT components (abstraction, algorithmic thinking, pattern recognition, decomposition) among 848 Estonian lower-secondary students. Overall, in this sample, girls outperformed boys, and ninth graders were overrepresented among High Achievers. However, these grade-level and gender differences should be interpreted cautiously because the ninth graders' subgroup was comparatively small and gender-imbalanced. The k-means analysis identified distinct patterns of performance across CT tasks. Because each task involved multiple CT components, the resulting clusters were interpreted as learner profiles rather than as indicators of specific CT subskills. Cluster membership was significantly associated with two single-item self-report indicators of IT-related intentions: High Achievers reported the greatest likelihood of IT studies/careers. Findings suggest CT competence (especially mastery across interdependent subskills) co-occurs with, but does not causally predict, higher self-reported intentions to pursue IT studies and careers.

Keywords: Computational thinking; Lower-secondary school students; Assessment; Cluster analysis; Intention.

1. Introduction

In today's technology-saturated world, where technological changes are tremendous and unpredictable, technology-related user competencies (both technical and conceptual) are essential for appropriately adapting to rapid changes. For example, the quality of Computing Education has been a key element in the European Commission's Digital Education Action Plan for the years 2021–2027 (European Commission, 2020), and Europe's Digital Decade (European Commission, 2021) policy programme sets the objective to have 20 million ICT specialists by 2030, keeping the gender convergence in focus.

At the same time, the question has arisen whether the knowledge related to the information technology (IT) field currently offered in the education system meets the new situation

* Corresponding author

and the needs of students. For example, there has been much discussion about the need to develop students' Computational Thinking (CT) skills, as the ability to use several computational elements supports a novel approach to solving emerging challenges not only in the field of technology but also in a broader context (Liu et al., 2024; Nouri et al., 2020; Shute et al., 2017; Yeni et al., 2024; Wing, 2006). Furthermore, CT has been framed as the core of STEM disciplines (Weintrop et al., 2016). Hence, the importance of developing CT has been articulated in various international development documents, e.g., ISTE Computational Thinking Competencies (International Society for Technology in Education, 2019), Computer Science Teachers Association's K–12 Standards (CSTA, 2017), European Commission DigComp 2.1: The Digital Competence Framework for Citizens (Carretero Gómez et al., 2017), as well as in national curriculums, e.g. in the United Kingdom (Department for Education, 2014), Australia (Australian Curriculum, Assessment and Reporting Authority, 2024), Portugal (Piedade et al., 2020), and Finland (Finnish National Board of Education, 2014). Altogether, at least 20 countries have defined CT as a crucial competence in the future in their curricula (Balanskat et al., 2017). In the Estonian national curriculum, CT has been set as one of two core elements of elective computer science courses, and it has been stated as a set of the following subskills: algorithmic thinking, pattern recognition, decomposition, and abstraction (Ministry of Education and Research, 2023).

Building on this widespread curricular commitment to CT, a practical question follows: how are students' CT task-response patterns associated with their self-reported intentions to pursue further studies in the IT field? This question is relevant because the persistent labor-market demand for IT professionals makes it important for education systems to understand how students come to see informatics-related pathways as possible options. As informatics is not a compulsory subject in several countries and it is not taught in all schools, but CT is a skill that can be developed across all subjects, the focus on teaching CT skills could solve the challenge of engaging and attracting students who do not have the opportunity to learn about IT in their informatics lessons.

However, it is currently unknown whether and how students' CT task-response patterns and their self-reported intention to continue their studies in the IT field are related, especially in contexts where prior engagement in informatics-related activities such as programming or robotics may vary across schools. It is also unclear how students respond to tasks that combine several CT components, rather than measuring separate CT subskills in isolation. Therefore, it is essential to continue research in this area.

The novelty of the present study lies in the combination of interpretation of CT components as interdependent rather than psychometrically isolated subskills, using cluster analysis to identify groups of students with similar learner profiles based on response patterns across multi-component CT tasks, and examining associations between cluster membership and students' self-reported intentions to pursue IT-related studies. This combination allows the study to move beyond comparisons of total CT scores and to examine how patterns of responses to multi-component CT tasks may be informative for differentiated support and guidance in informatics education.

2. Theoretical Framework

CT, one's ability to use the thought processes for formulating challenges and their solutions in a form that any other information-processing agent can also carry out the solution, has been regarded as one of the most critical skills in today's technology-based world (Angeli & Giannakos, 2020; Liu et al., 2024; Nouri et al., 2020; Shute et al., 2017; Tariq et al., 2025; Wing, 2011). According to Wing (2006), it should be taught to students from an early age alongside reading, writing, and arithmetic. The concept has its origins in the "theory of computation" from the 1960s but was systematically described by Wing (2006) as a competence for "solving problems, designing systems, and understanding human behaviour, by drawing on the concepts fundamental to computer science." In general, CT has been defined as a combination of skills inherent to computer science, problem-solving competence, and creativity, but there is still no consensus on the skills that should be involved, and models of CT vary in different studies (Barr & Stephenson, 2011; Bounou et al., 2023; Stella et al., 2021).

Although the development and importance of CT skills were initially seen primarily in the context of computer science, their significance soon began to be highlighted more broadly in education and cognitive sciences (Aydeniz, 2018). It has also been acknowledged that the capability to develop students' CT skills is not unique to computer sciences, but many disciplines can create it (Adeoye & Prastikawati, 2024).

CT is a multidimensional concept, and there is still no definitive consensus on the skills that complete the model (Brennan & Resnick, 2012). Wing (2006) developed a six-skill model in which problem formulation, abstraction, problem reformulation, decomposition, automation, and systematic testing were identified as CT components. Selby and Woollard (2013) reviewed the literature and described CT as cognitive processes associated with, but not limited to, problem-solving. Their model included abstractions, decomposition, algorithmic design, evaluation, and generalization. Based on the literature, Shute et al. (2017) defined skills such as decomposition, abstraction, algorithm design, debugging, iteration, and generalization. International Society for Technology in Education (2019) set a model where decomposition, gathering and analyzing data, abstraction, and algorithm design are the skills to develop to understand and solve problems in ways that leverage the power of technological methods. Palts and Pedaste (2020) reviewed 65 articles and designed a CT model where skills were categorized into three sequenced stages: defining the problem, solving the problem, and analyzing the solution. The European Commission's study "Reviewing Computational Thinking in Compulsory Education" (Bocconi et al., 2022) stated abstraction, algorithmic thinking, automation, decomposition, debugging, and generalization as core skills of CT.

All models under the survey included abstraction and decomposition (Bocconi et al., 2022; International Society for Technology in Education, 2019; Palts & Pedaste, 2020; Selby & Woollard, 2013; Shute et al., 2017; Wing, 2006), a majority of models included algorithmic design (Bocconi et al., 2022; International Society for Technology in Education, 2019; Palts & Pedaste, 2020; Selby & Woollard, 2013; Shute et al., 2017) and generalization (Bocconi et al., 2022; Palts & Pedaste, 2020; Selby & Woollard, 2013; Shute et al., 2017).

As can be seen, several models describe primary skills (path to solution) similarly but differ in aspects concerning activities after finding the solution to the task. For example, Wing (2006) and Palts and Pedaste (2020) describe systematic testing, also evaluation (Palts & Pedaste, 2020; Selby & Woollard, 2013) and debugging (Bocconi et al., 2022; Shute et al., 2017) have been phrased. Skills like iteration and parallelization have been brought out in the model described by Palts and Pedaste (2020), and the former has also been described by Shute et al. (2017).

To sum up, it is possible to conclude that algorithmic thinking, pattern recognition, decomposition, and abstraction are these four subskills that capture the core essence of CT frameworks: breaking problems down into small parts to make them easier to solve (decomposition), recognizing a pattern or similar characteristics to help break down the problem, build a construct as a path for the solution, and find a set of patterns or similar characteristics that can be generalized (pattern recognition), taking away or removing characteristics from something to reduce it to a set of essential characteristics (abstraction), and practice of writing a step-by-step sequence of instructions for a solution or process (algorithmic thinking). Large-scale policy frameworks, such as Google Research, operate CT around the same four pillars (Google Research, n.d.).

2.1 Possible Ways to Evaluate CT

Measuring students' learning outcomes is necessary to assess the effectiveness of learning and the suitability of teaching methods. Without a focus on assessment, it is unlikely that CT will be effectively integrated into any curriculum (Román-González et al., 2017). Therefore, it is necessary to contribute to creating a system of diverse assessment approaches, combining different tools and instruments to focus on various aspects of CT skills (Grover, 2021).

Over the years, researchers have used several approaches to measure students' CT skills, and several papers have addressed the development of valid and reliable instruments. Palts (2021) has suggested that, based on the tools used, scenarios for developing and assessing CT skills can be grouped into three main categories: computer-based activities, such as gaming and project creation; robotics and tinkering activities; and unplugged activities. To evaluate students' CT skills, tests have often been used (Bounou et al., 2023; Gugemos, 2021; Werner et al., 2012) and selected-response items are mainly used in this case (Cutumisu et al., 2019), but also open-ended programming assignments have been in use (Grover, 2017). In other cases, tests have been combined with other research instruments to provide a broader view of assessment (Li et al., 2021a). With younger students, observations and interviews have been conducted, while broader projects have served as assessment instruments in middle school (Grover, 2017). In addition, gaming environments and simulations have been created to assess aspects of CT (Grover & Pea, 2013), and pair programming has served as a tool to evaluate students' performance in that field (Werner et al., 2012).

2.1.1 *Bebras tasks as an assessment option for computational thinking*

In the survey, the Bebras International Challenge on Informatics and Computational Thinking (Bebras) tasks are used to assess students' CT skills. Bebras offers short, age-appropriate tasks that evaluate CT skills without requiring prior programming experience (Bebras, n.d.). Bebras tasks are usually presented as brief problem scenarios supported by visual context and are most often delivered in a selection-response format, making them feasible for large-scale administration in regular school settings (Nakić et al., 2025). Compared with other CT assessment options, Bebras' tasks offer several practical advantages for the present study context. First, they do not depend on the school's curriculum, as it is possible to successfully solve CT tasks without prior programming or robotics experience: rather than coding, these tasks focus on reasoning and problem-solving. Second, their short format supports limited testing time and classroom logistics. Third, the variety of available tasks makes it possible to cover multiple CT subskills and difficulty levels (Dagienė & Sentance, 2016). Fourth, in Estonia, Bebras (Kobras, n.d.) tasks have been used nationally in Bebras competitions over multiple years and are selected, translated, and reviewed by the national Bebras organisation, which supports linguistic and cultural appropriateness and reduces the risk associated with ad hoc translation of new instruments.

2.2 *Previous Studies on Differences in CT Skills and Their Associations with Students' IT Study Intentions*

Previous studies have been trying to find connections between CT skills and several aspects of students' competencies to find possibilities to foster the development of CT skills. For example, as CT skills are something which could be taught in a variety of school subjects, research to find ways to predict relations between success in some subjects and CT skills has been conducted. There are findings that refer to the positive impact of achievement in STEM subjects, especially in math, on CT performance (Bounou et al., 2023; Chan et al., 2020; Chongo et al., 2020; Şişman et al., 2021). Interestingly, it was found that achievement in IT affected CT performance less than achievement in math (Şişman et al., 2021). Not only the achievements but also the positive attitude towards math have been proven to have a positive impact on CT skills (Didem, 2020).

Several researchers have assessed students' programming skills, considering these skills to represent students' levels of CT (Bounou et al., 2023). Some studies have found that students with programming experience have significantly better skills at solving CT tasks (Cachero et al., 2020; Li et al., 2021a). It has also been stated that students with self-perceived programming proficiency (Li et al., 2021b) or a positive attitude to programming (Sun et al., 2022) have significantly better skills in solving CT tasks. At the same time, the relationship between CT and IT skills is often viewed as bidirectional (Dohn et al., 2022). On the one hand, developing CT skills can enhance a student's ability to acquire IT skills; on the other, having a strong foundation in IT skills can facilitate the understanding and application of CT concepts (Cui et al., 2023; Durak et al., 2017). This reciprocal relation-

ship suggests that integrating CT into IT education may improve outcomes in both areas.

Some studies have broadened the background for evaluating CT skills in connection with problem-solving, critical thinking, creativity, and communication (Corrales-Álvarez et al., 2024). The studies suggest that CT skills are most likely connected to students' problem-solving and critical thinking skills, showing positive correlations between CT and both mentioned aspects (Kocak et al., 2021; Wu et al., 2024). However, as CT is still a construct in which precise elements have not been finally agreed upon, its assessment also remains an unresolved issue (Román-González et al., 2018).

2.2.1 Previous Studies on Differences in CT Skills

Research into age- and schooling-level differences in CT skills has shown that students' ability to solve CT tasks may improve with age and schooling level, yet similar patterns of errors appear across different age groups (Mannila et al., 2016; Strawhacker & Bers, 2019; Teng & Chung, 2025). In general, differences in performance by age are particularly clear in tasks that require abstraction and decomposition subskills (Teng & Chung, 2025). At the same time, some research finds that while age differences in these mentioned subskills can be found in primary-school tasks, they disappear from other CT subskills and at later stages of schooling (Mertens & Colunga, 2025). To sum up, the better performance of older students can come from general abilities, such as higher cognitive skills and better working memory (Román-González et al., 2017), but it can also be affected by prior programming experience, which is naturally longer for older students (Li et al., 2021b).

Multiple studies have examined gender differences in CT skills among school-aged students. Some research suggests that boys often outperform girls in overall CT performance, with significant differences observed in both CT skills and perceptions of ability, particularly in secondary school settings (Chan et al., 2020; Ma et al., 2021; Román-González et al., 2017; Şişman et al., 2021; Sun et al., 2023). However, some research indicates that with targeted interventions, such as gamified programming or educational robotics, boys and girls can achieve similar levels of CT skills by the end of the learning period, although girls may require more time or specific support to reach parity (Atmatzidou & Demetriadis, 2016; Sun & Liu, 2023) and they tend to pay more attention to group dynamics and collaboration during CT activities, while boys may focus more on operational aspects like building and coding (Ardito et al., 2020). Thus, boys' better results may not indicate higher skills, but rather better suitability of the activities used for measurement for them. There are also studies where significant differences between boys and girls have not been found (Katai et al., 2021; Li et al., 2021b).

To sum up, while boys sometimes initially outperform girls in CT skills, well-designed educational interventions and inclusive practices could help to bridge these gaps. Thus, to keep gender convergence in focus, more diverse learning methods could help provide equal opportunities for all to improve their CT-related competencies.

2.2.2 Associations Between CT Skills and Students' Intentions to Continue Their Studies in the IT Field

As CT continues to be integrated into curricula worldwide, establishing whether and how CT competencies are associated with students' self-reported interest in technology-related study pathways has become increasingly urgent for educational policy, targeted interventions, workforce development, and for understanding how technical skill development may be connected with students' educational intentions (Bocconi et al., 2022). Based on the research, it can be said that direct studies specifically examining whether CT skills are associated with students' intentions to continue in IT fields remain limited.

Existing studies have mainly focused on related factors that may shape future educational and career choices, such as self-efficacy, motivation, and broader STEM career interest, rather than directly testing IT-specific intention (Hava et al., 2021; Angeli & Giannakos, 2020; Chen et al., 2021; Zheng et al., 2024; Wang et al., 2024). Therefore, the direct link between CT skills and intention to pursue further studies or careers in IT remains insufficiently explored. This gap is crucial to note, given the widespread integration of CT into educational curricula and the growing demand for IT professionals globally. When investigating the question more broadly, studies suggest that students who develop CT competencies may be better positioned to explore and commit to STEAM and assumingly also to IT-related career paths (Hava et al., 2021), as CT competence is positively correlated with students' self-efficacy and motivation in computer science, and these aspects may, in turn, be associated with career-related intentions in that field (Angeli & Giannakos, 2020; Chen et al., 2021; Zheng et al., 2024; Wang et al., 2024). These findings suggest that although CT may be linked to broader motivational and career-related factors, more targeted research is needed to clarify whether and how CT skills are associated with students' self-reported intentions to pursue further studies and careers in IT.

2.3 Aim of the Study and Research Questions

Based on previous studies, it is possible to assume that students' intention to continue their studies in the IT field may be associated with prior exposure to IT and the students' perception that they can cope in that field (Angeli & Giannakos, 2020; Bandura, 1997; Chen et al., 2021; Zheng et al., 2024; Wang et al., 2024). However, it cannot be expected that all schools offer the same amount or type of informatics education. For example, in Estonia, informatics is included in the national curriculum as an elective subject (Ministry of Education and Research, 2023) and subjects belonging to the field of informatics are taught in the majority of schools, but there is no universally implemented concept of teaching, and often the content is limited based on the teacher's expertise or the need to focus on broader digital skills and not IT skills (Luik et al., 2023). This, in turn, leads to differences or, in some cases, even a lack of students' informatics exposure. Therefore, it is necessary to find alternative ways to support students' willingness to consider studying in the IT field.

In the context of the updated Estonian curriculum (Ministry of Education and Research,

2023), whose flexibility and emphasis on competencies and transversal skills grant teachers considerable autonomy in shaping the learning process, the development of CT skills can be meaningfully embedded across a range of subjects through diverse instructional approaches, rather than being confined to the content of informatics courses. Therefore, it could be one possible way to support students' self-efficacy and interest in the IT field. However, it is still unclear whether there is a link between students' CT skills and their intention to link their future to IT, as research in this field is limited.

The aim of the study is to identify patterns in students' responses to CT tasks and to examine how these patterns are associated with students' self-reported intentions to continue their studies in the IT field.

The study was based on the following research questions:

RQ1: What similarities and differences are observed in this sample between boys and girls, and between students from different grades, in solving CT tasks?

RQ2: What patterns of correct responses across the 12 CT tasks characterize the identified clusters?

RQ3: Is there a statistically significant relationship between students' belonging to a specific CT task-response cluster and their self-reported intention to pursue further studies in the IT field?

3. Methodology

3.1. Sample

A total of 1088 students responded to the quantitative survey. Incomplete responses ($n = 218$) were excluded from the analysis, and any responses for which the school's name could not be identified were also excluded from the analysis. The final sample of the quantitative survey included 848 students from lower secondary levels: 7th grade (13-year-olds), 8th grade (14-year-olds), and 9th grade (15-year-olds). The exact division of the sample is shown in Table 1. The students were chosen from 30 schools in Estonia, which is about 10% of all schools offering lower-secondary education (Haridussilm, 2023). The selection included the correct proportion of schools from rural and urban areas. The type of school was also considered, and students from middle schools, K–12 schools, and public and private schools participated.

Table 1. Participants in the survey.

Grade	Participants n (%)	Females n (%)	Males n (%)
7th	376 (44.3%)	196 (52.1%)	180 (47.9%)
8th	399 (47.1%)	199 (49.9%)	200 (50.1%)
9th	73 (8.6%)	55 (75.3%)	18 (24.7%)
Total	848 (100%)	450 (53.1%)	398 (46.9%)

Altogether, 461 (54%) students were from urban areas and 387 (46%) from rural areas. According to the statistics (Haridussilm, 2023), the division of the sample by gender and urban/rural status matched the actual situation in Estonia.

For the pilot phase, 25 students (57% females and 43% males) from two class sets were selected, and a qualitative survey using the think-aloud method was conducted with them.

3.3.1 Instrument and Procedure

Since the majority of analysed studies (Cutumisu et al., 2019; Katai et al., 2021; Li et al., 2021a) have used a test to evaluate students' CT skills and to define various skills in the CT concept, the test was also used in this study. The purpose was to identify distinct groups of students based on patterns of correct and incorrect responses across CT tasks, using tasks aligned with the four CT components stated in the Estonian national curriculum (Ministry of Education and Research, 2023): abstraction, algorithmic thinking, pattern recognition, and decomposition, rather than to produce psychometrically isolated subskill scores. According to the literature review conducted by Piedade et al. (2020), the factors were defined as seen in Table 2.

Table 2. Description of CT subskills.

Subskill	Description
Abstraction	The process of taking away or removing characteristics from something to reduce it to a set of essential characteristics.
Algorithmic thinking	The practice of writing a step-by-step sequence of instructions for a solution or process.
Decomposition	Breaking problems down into smaller parts to make them easier to solve.
Pattern recognition	Recognizing a pattern or similar characteristics to help break down the problem, build a construct as a path for the solution, and find a set of patterns or similar characteristics that can be generalized.

The composition of the test consisted of several steps. First, potentially suitable tasks were selected for use in the test. The Bebras international informatics challenge tasks were used to select tasks, as the tasks from this competition follow the CT model and skills chosen for the study and are suitable for the age group of the students in the sample (Dagienė & Sentance, 2016). Tasks were selected from challenges between 2017 and 2021, and initial categorization was based on the skills phrased in the tasks by their creators. Three tasks were selected for each subskill to ensure enough tasks were included in the final test version after the experts assessed them.

The suitability of the tasks for the age group was assessed by two teachers, co-authors of the study, and four students from the age group. Based on their consensus, some tasks were excluded due to their simplicity, complexity, or potentially confusing task descriptions. The final survey consisted of 12 tasks initially selected to cover the four CT components. The tasks were divided equally (three tasks initially associated with each subskill). The

aim was to offer tasks of varying difficulty levels in each subskill. Each task consisted of a problem description with a visual and a question; 11 were multiple-choice items, and 1 was an open-ended task. A grading system was set up where the correct answer received 1 point and the wrong one 0 points.

Then, the pilot survey was conducted to evaluate clarity and timing and to validate how tasks could be mapped to CT subskills. For piloting, two sessions were conducted. The first pilot involved 14 students, and the second, 11. Both groups participated in a single 45-minute session facilitated by the author. The Concurrent Think-Aloud (CTA) protocol was used: students verbalized their thoughts while completing the tasks. During the sessions, the researcher acted as an observer and took detailed written notes on students' verbalizations and observable task-solving behaviour. The aim was to grasp the idea of students' thinking process, so it would be possible to find whether the steps used by solvers are in accordance with the description of steps characteristic of subskills described in Table 3. The pilot material was analyzed using qualitative protocol analysis. Coding focused on identifying evidence of CT subskills in students' problem-solving processes based on the content of the think-aloud protocols (i.e., which subskills could reasonably be attributed given the students' verbalized reasoning and behaviour). The preliminary coding results were discussed with the co-authors, and final decisions regarding the task and subskill assignments were made using a consensus-based qualitative validation procedure. This process supported the final mapping presented in Table 3 and supported the interpretation that individual tasks typically evaluate multiple CT subskills rather than a single subskill.

Table 3. The tasks in the test and the subskills they evaluate.

Task	Title of the task*	CT subskills**
Q1	School lunch	AT, PR, DEC
Q2	Image layers	AT, PR, DEC
Q3	Cross-stitch embroidery	AT, PR, DEC
Q4	Animal competition	AT, PR, DEC, AB
Q5	Forest pictures	AT, PR
Q6	Animation	AT, PR, AB
Q7	Swinging	DEC, AB
Q8	Geocaching	PR, DEC, AB
Q9	Walnut animals	AT, DEC, AB
Q10	Crane operating	AT, DEC
Q11	Beaver games	AT, PR, DEC, AB
Q12	Mushrooms	AT, DEC, AB

* Original title from the Bebras competition

**AT – algorithmic thinking, PR – pattern recognition, DEC – decomposition, AB – abstraction

The analysis of the test results and the think-aloud method indicated that categorizing tasks as measuring only one isolated CT subskill is not reasonable or sufficient. Consequently, at

this stage, the authors reassessed the CT subskill knowledge that each task measured, utilizing the classification presented in Table 2, which outlines subskills' descriptions. During this process, the researchers assigned the subskills they deemed appropriate to each task individually. Then, they reached a consensus through discussion. This led to a new mapping of the tasks. Ultimately, it was found that none of the tasks assessed only a single CT subskill: for all tasks, at least two subskills were identified. In two tasks, the authors discovered that measuring all four subskills was possible. The tasks and their subskills can be found in Table 3. The task-to-subskill mapping is intended to characterize the CT components involved in solving each task. It does not imply that the tasks function as pure indicators of individual CT subskills. As all tasks require combinations of CT components, the instrument is better conceptualized as a CT task set assessing integrated computational thinking performance rather than as a four-factor subskill instrument.

The pilot survey was conducted in April 2023 to evaluate clarity and timing and to examine how the tasks could be mapped to CT subskills. After the piloting, the test was carried out. The data was collected from April to June 2023. Students filled out the electronic questionnaire form in LimeSurvey. The test ended with background data (gender and grade level). Assent was obtained immediately before the data collection (tick-box assent in the online questionnaire, verbal assent for think-aloud sessions). The assent procedure explained voluntary participation, the right to withdraw at any time without consequences, and that responses would not affect grades or school evaluation. Since the participants were underage, informed consent was obtained from the parents or legal guardians of all participants involved in the survey, ensuring they were fully aware of the study's nature and purpose. Data was collected and processed in accordance with GDPR principles (lawfulness, data minimization, purpose limitation). Students' responses were stored pseudonymized. The study received approval from the University of Tartu Centre for Ethics (approval number 385/T-22), affirming that all procedures followed ethical standards. All data collected from the survey is stored according to the requirements, and only the research team has access to the key file. Data were stored on the University's server, with access restricted to named researchers. The data will be used solely for this study, with any publications resulting from this research anonymizing identifying information to prevent participant identification.

Basic psychometric properties were calculated for the 12-item CT task set. Because all tasks were scored dichotomously (0 = incorrect, 1 = correct), internal consistency was estimated using the Kuder–Richardson Formula 20 (KR-20). The analysis indicated adequate internal consistency for the exploratory use of the test as a holistic task set (KR-20 = 0.775), as values around .70 are often considered acceptable for internal consistency, particularly in exploratory research (Taber, 2018; Edelsbrunner et al., 2025). Item difficulty was reported as the proportion of students who answered each task correctly and ranged from 0.298 to 0.916. Item discrimination was examined using corrected item-total correlations, calculated as the correlation between each task score and the sum score of the remaining 11 tasks. These correlations ranged from 0.253 to 0.548, indicating that all items showed positive discrimination with the remaining total score.

3.1 Data Analysis

IBM SPSS Statistics (version 30.0) was used for statistical data analysis. In the preliminary analysis, incomplete responses ($n = 218$) were removed. Item-level descriptive statistics (the percentage of correct responses, mean and standard deviation) were used for all tasks coded dichotomously (1 = correct, 0 = incorrect).

To identify distinct patterns in students' CT performance based on item-level responses, a k-means cluster analysis was conducted using Euclidean distance and a maximum of 99 iterations. Because all 12 task variables were dichotomous and on the same metric (0 = incorrect, 1 = correct), no further standardization was required. In the present study, k-means was used as an exploratory partitioning tool to summarize recurring CT response patterns rather than to infer latent categories. As k-means cluster analysis is performed with different numbers of clusters to find a solution that is the most meaningful in the research context (Jain, 2010), we tested the cluster models with three, four, and five clusters. A four-cluster solution was selected as it showed the most interpretable and clearly differentiated task-response patterns with meaningful cluster sizes. Cluster characteristics were described using within-cluster percentages and descriptive statistics. Because cluster membership was derived from the same 12 task responses, the task-level patterns were used only to describe the patterns. We did not treat post-clustering comparisons of these defining variables as independent inferential tests. The clustering was conducted as an exploratory and descriptive analysis: the selected solution should therefore be understood as a data-driven summary of task-response patterns in the present sample, rather than as evidence of definitive, fixed, or generalizable learner categories.

Associations between cluster membership (cluster) and categorical variables (gender, grade level) were examined using chi-square tests. When the overall chi-square test was significant, follow-up pairwise chi-square comparisons were conducted to identify which cluster pairs differed, with Bonferroni-adjusted p-values. Differences in overall CT test scores across grade levels and gender were examined using one-way ANOVA with Bonferroni-adjusted post-hoc comparisons. Differences in students' IT-related intention ratings (0–100%) across clusters were also examined using one-way ANOVA with Bonferroni-adjusted post-hoc comparisons. In addition to p-values, 95% confidence intervals were reported for group means and, where relevant, for mean differences to indicate the precision of the estimates. For ANOVA results, effect sizes were reported as eta squared (η^2).

4. Results

The overall analysis of the test results showed that the level of complexity of the tasks was in accordance with the age group, as we can see the normal distribution of responses (Figure 1). 31 respondents (3.7%) received the maximum (12) points, 4 respondents (0.5%) did not answer correctly to any of the tasks, receiving 0 points from the test.

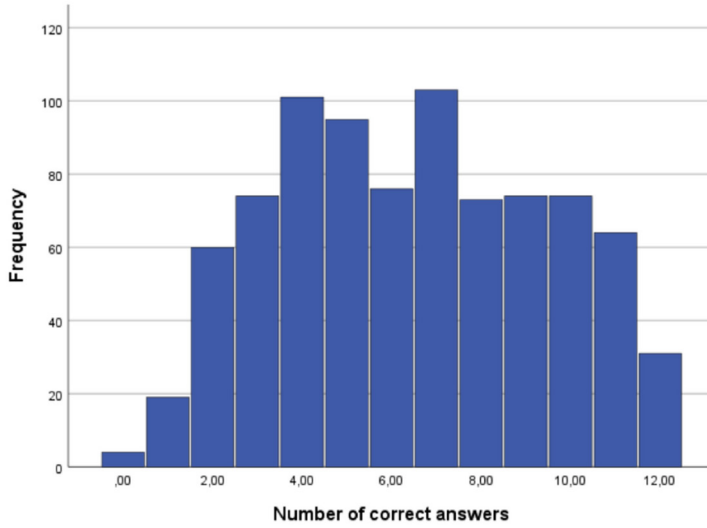


Figure 1. The distribution of respondents by results.

Considering all the tasks, the mean score students obtained was 6.47 ($SD = 2.98$). In general, task 2, where students needed to decide the order of placing layers in an image, was the easiest (mean 0.92, $SD 0.28$). According to our task mapping, algorithmic thinking, pattern recognition, and decomposition were the subskills involved in this task. Several tasks with lower results (tasks 4, 12, and 9) included abstraction, but they also involved other CT components. Therefore, these results should be interpreted cautiously as weaker performance on multi-component tasks that included abstraction, not as evidence of an isolated abstraction deficit. There was one exception, task 10 (Crane operator), that did not include abstraction, but where students received relatively low results. This task required the ability to remember a noticeably higher number of steps than any other task to solve the task correctly, and minor inattention could lead to an incorrect answer.

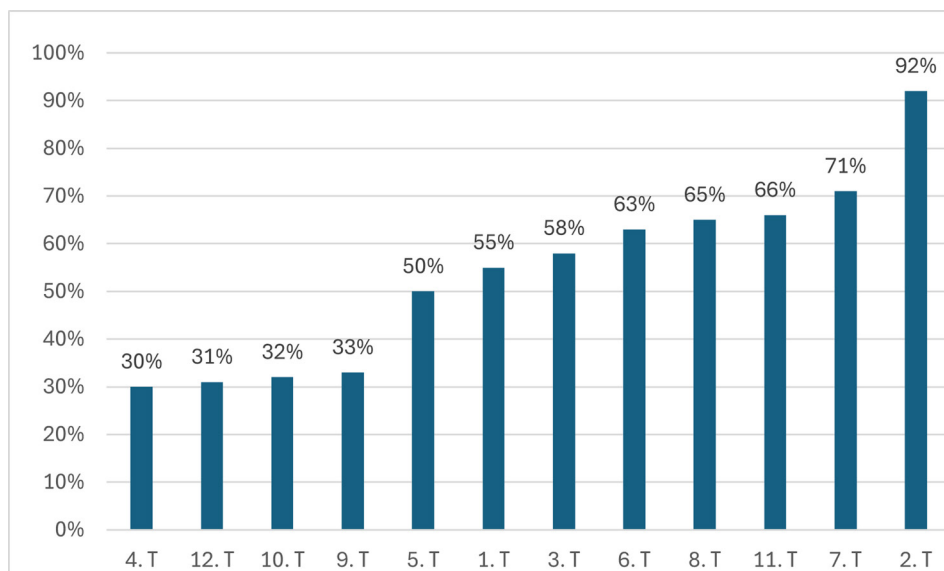


Figure 2. The percentage of students with the correct answer.

* T = task

Regarding overall CT scores by grade level, a one-way ANOVA indicated a statistically significant difference ($F(2, 845) = 30.58, p < 0.05, \eta^2 = .067$). Bonferroni-adjusted post hoc tests showed that 9th graders outperformed 8th ($p < 0.05$) and 7th ($p < 0.05$) graders with their overall test results, with mean scores of 8.99 (SD = 2.73, 95% CI [8.35, 9.63]) for 9th graders, 6.25 (SD = 2.89, 95% CI [5.97, 6.53]) for 8th graders, and 6.21 (SD = 2.90, 95% CI [5.92, 6.50]) for 7th graders. Also, none of the 9th graders received 0 points from the test. The difference between 8th graders and 7th graders was not statistically significant ($p > 0.05$).

When analyzing the tasks one by one, the statistically significant difference ($p < 0.05$) between grades (9th graders performing better than younger grades) came out in all tasks, except in tasks called “Image layers” and “Swinging”. These two tasks also received the highest scores overall (mean score 0.92 and SD = 0.28 for task “Image layers,” and mean score 0.71 and SD = 0.45 for task “Swinging”).

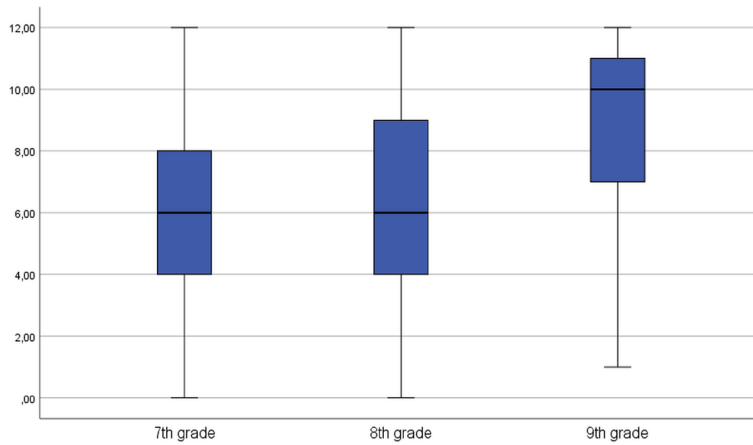


Figure 3. The test results by grades.

When comparing boys and girls, a one-way ANOVA showed a statistically significant but small difference ($F(1, 846) = 6.37, p < 0.05, \eta^2 = .007$) in overall CT scores: girls outperformed boys. Group means and 95% confidence intervals were as follows: girls, $M = 6.7$, $SD = 2.9$, 95% CI [6.43, 6.98]; boys, $M = 6.2$, $SD = 3.0$, 95% CI [5.9, 6.5]. When analysing the tasks one by one, the statistically significant difference ($p < 0.05$) between genders (in favor of girls) came out in five tasks (2, 5, 6, 11, and 12), but these tasks did not have any connecting feature distinguishing them from tasks that were not statistically different.

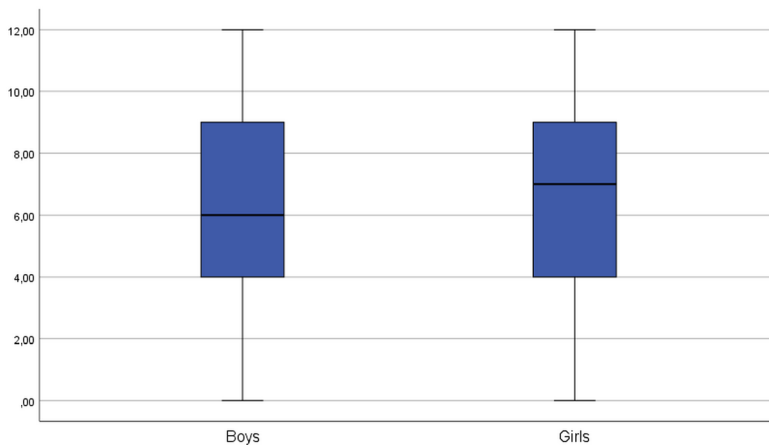


Figure 4. The test results by gender.

Cluster analysis was conducted to identify distinct item-level CT task-response patterns that could not be fully captured through descriptive statistics alone. The k-means method was applied, and several potential solutions were compared. Based on these criteria, a four-cluster solution was selected, as it yielded the most interpretable and distinctive patterns in the data. The identified clusters offer a framework for examining variation in CT task-response patterns, not as indicators of separate CT subskills in isolation (Figure 5).

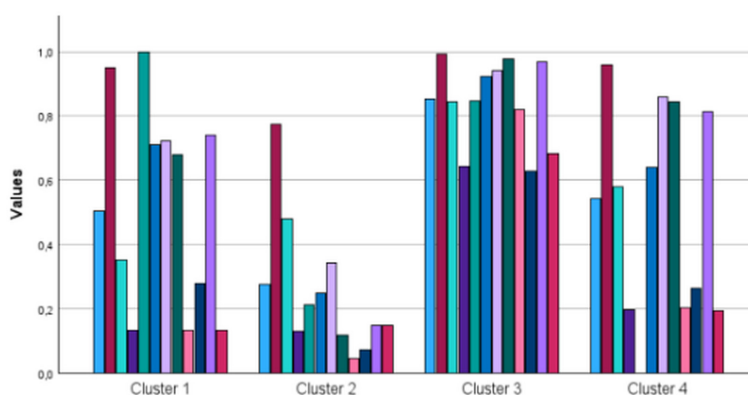


Figure 5. Four CT task-response profiles identified by k-means clustering.*

* The x-axis shows the four clusters, and the y-axis shows the proportion of students within a cluster who answered a task correctly (0 = no students in the cluster answered correctly; 1 = all students in the cluster answered correctly). Within each cluster, the bars from left to right correspond to Q1–Q12 as listed in Table 3.

The following cluster descriptions summarize the response patterns that formed the four-cluster solution. Because the same 12 task responses were used to derive cluster membership, these task-level patterns are descriptive characteristics of the patterns rather than independent statistical tests of differences between clusters.

The first cluster comprised 20% of respondents. This profile showed comparatively lower proportions of correct responses on tasks 3, 7, and 8, and comparatively higher proportions on tasks 2 and 5. The lowest proportions of correct responses within this profile were observed on tasks 4, 9, 10, and 12. Tasks 2 and 5 involved pattern recognition and algorithmic thinking, whereas several of the tasks with lower correct-response proportions involved decomposition and/or abstraction in combination with other CT components. We therefore labeled this profile Pattern Recognisers. The label is descriptive shorthand for the observed response pattern and does not indicate a distinct or isolated CT subskill.

The second cluster comprised 28% of respondents. This profile showed comparatively low proportions of correct responses across most tasks, particularly tasks 1, 2, 6, 7, 8, 11, and 12. Within this profile, task 2 represented a relative strength, whereas tasks 4, 8, and 9 showed the lowest correct-response proportions. As this profile reflected broadly lower performance across the task set, we labeled it Low Achievers. This label describes the observed task-response pattern in the present sample and should not be interpreted as a fixed learner category.

The third cluster comprised 29% of respondents. This profile showed comparatively high proportions of correct responses on most tasks, especially tasks 1, 3, 4, 6, 8, 9, 10, 11, and 12. Within the profile, tasks 4, 10, and 12 had relatively lower correct-response proportions than the other tasks, although the profile still showed broadly strong performance across the task set. We labeled this profile **High Achievers** to describe its comparatively strong and broad task-response pattern. The label does not provide psychometric evidence of mastery of each CT component in isolation.

The fourth cluster comprised 23% of respondents. This profile showed comparatively higher proportions of correct responses on tasks 7, 8, and 11, whereas task 2 was the relative strength within the profile. Lower correct-response proportions were observed on tasks 5, 4, 9, 10, and 12. According to the task mapping, tasks 7, 8, and 11 involved decomposition and abstraction in combination with other CT components, while task 5 involved pattern recognition and algorithmic thinking. We labeled this profile **Structured Thinkers** as descriptive shorthand for this task-response pattern. It should not be interpreted as evidence of an isolated algorithmic-thinking deficit or of separate mastery of decomposition and abstraction.

A chi-square test revealed a statistically significant association between gender and cluster membership ($\chi^2 = 9.107$, $p < 0.05$, Cramér's $V = 0.104$). Pairwise comparisons showed significant differences between Low Achievers and High Achievers ($p < 0.05$), indicating that the Low Achievers cluster included more boys than the High Achievers cluster. No significant gender differences were found between Pattern Recognisers and High Achievers or Structured Thinkers, nor between Low Achievers and Structured Thinkers. A chi-square test also revealed a statistically significant association between cluster membership and students' grade levels ($\chi^2 = 53.68$, $p < 0.05$, Cramér's $V = .178$). Pairwise comparisons indicated that High Achievers stood out with a significantly higher proportion of 9th-grade students, differing from all other clusters ($p < 0.05$). No significant grade-level differences were found between Pattern Recognisers and Low Achievers or between Low Achievers and Structured Thinkers. Because 9th grade was comparatively small and gender-imbalanced, and because prior informatics exposure or extracurricular IT experience was not measured, these cluster-level grade and gender associations should be interpreted cautiously.

When examining differences in students' self-reported intention to continue studies in the IT field (question: "How likely is it, in percent (0–100%), that you will pursue studies in an IT-related field in the future?") and in students' self-reported intention to link their future with the IT field (question: "How likely is it, in percent (0–100%), that you will pursue a future in IT?") across the four CT clusters, the one-way ANOVA test revealed statistically significant between-cluster differences for self-reported intention to pursue IT-related studies ($F(3, 844) = 9.458$, $p < .05$, $\eta^2 = .033$) and for self-reported intention to pursue a future in IT ($F(3, 844) = 13.040$, $p < .05$, $\eta^2 = .044$). Descriptive means (mean intention ratings) with 95% confidence intervals are presented in Table 4.

Table 4. Differences in students' self-reported IT-related intention ratings across clusters.

	Mean and 95% CI				F	P	η^2
	Pattern Recognisers	Low Achievers	High Achievers	Structured Thinkers			
Q1*	24 [19.9, 27.7]	23 [19.2, 26.1]	34 [30.9, 37.9]	26 [22.5, 29.5]	9.458	< 0.05	.033
Q2**	27 [23.0, 31.7]	25 [21.8, 29.9]	40 [36.3, 43.8]	29 [24.8, 32.9]	13.040	< 0.05	.044

* How likely is it, in percent (0–100%), that you will pursue studies in an IT-related field in the future?

** How likely is it, in percent (0–100%), that you will pursue a future in IT?

Post-hoc pairwise comparisons (Bonferroni-adjusted) showed significantly higher mean self-reported intention ratings for both “intend to pursue their studies” and “pursue a future in IT” among High Achievers than in any other cluster ($p < 0.05$). No statistically significant differences were observed between other clusters.

5. Discussion

This study aimed to explore the possible associations between students' CT task-response patterns and their self-reported intentions to continue their studies in the IT field. The cluster analysis was used to identify distinct holistic CT task-response patterns among lower-secondary school students and to identify differences and similarities among those patterns. Because each task involved multiple CT components (abstraction, decomposition, algorithmic thinking, and pattern recognition), the clusters should be interpreted as patterns of students' performance across CT tasks rather than as representations of individual CT subskills. From an educational perspective, these patterns provide useful information about the types of tasks students tend to solve successfully and those that may require additional support. Such information may help teachers tailor feedback, learning activities, and guidance to students' needs. Furthermore, the association between cluster membership and students' self-reported IT study intentions suggests that broader patterns of CT task performance are related to students' educational interests. However, these findings should be interpreted as correlational and do not imply that specific CT subskills directly influence students' intentions to pursue IT studies.

The first research question focused on similarities and differences between boys and girls and between grade levels in solving CT tasks. The overall results reflected moderate test difficulty together with substantial variation in performance across tasks, indicating that some task types were consistently more challenging than others. From a classroom perspective, this suggests that short Bebras-based CT tasks can serve not only as assessment instruments but also as sources of formative information about students' performance patterns. By examining which tasks students solve successfully and which present challenges, teachers may identify areas where additional support or practice is needed. Such information can be used to design follow-up activities that address the cognitive demands reflected in the tasks and to provide more targeted feedback and guidance.

In general, 9th graders received higher results than 8th and 7th graders in this sample, and the effect size of this difference was medium, suggesting that grade level had a meaningful association with students' overall CT performance. It remains unclear whether the better results in 9th grade are due to age, more prolonged exposure to programming (Li et al., 2021b), or stronger overall cognitive abilities, as also stated in several previous studies (Román-González et al., 2017). This is particularly relevant because some studies have found no statistically significant relationship between age and CT skills among university students, suggesting that age alone may not account for differences in CT performance (Atman Uslu, 2023). Therefore, the observed grade-level differences are likely to reflect a combination of developmental, educational, and contextual factors. Also, as the respondents' group from 9th grade was comparatively small and gender-imbalanced, future research is needed to disentangle the relative contributions of age, educational experiences, and prior exposure to programming.

Overall, girls achieved statistically significantly better results than boys in this sample. However, the effect size was very small, indicating that although the difference reached statistical significance, its practical magnitude was limited. In general, previous research suggests that boys and girls do not necessarily differ in CT knowledge, and observed differences may partly depend on the style, content, and format of the assessment tasks. In the present study, girls tended to outperform boys particularly on tasks that required combining information from visual representations and written descriptions. One possible explanation is that such tasks place greater demands on reading comprehension and the integration of textual and visual information. This interpretation is consistent with findings reported by Sousa and Tomlinson (2011), who suggest that developmental differences in language-related skills may influence performance on certain types of learning tasks. However, the present study was not designed to test these mechanisms directly. Therefore, the observed difference may reflect a combination of task characteristics, prior experiences, and contextual factors rather than a stable difference in CT ability between girls and boys. The finding nevertheless highlights the importance of using diverse CT task formats and instructional approaches so that assessment and learning opportunities are accessible to students with different strengths and problem-solving preferences.

The second research question asked what patterns of correct responses across the 12 CT tasks characterized the identified clusters. The four-cluster solution summarized distinct patterns of correct responses across tasks that integrated multiple CT components. Pattern Recognisers showed comparatively higher proportions of correct responses on tasks 2 and 5 and lower proportions on several tasks involving decomposition and/or abstraction. Low Achievers showed comparatively low correct-response proportions across most tasks, with task 2 being a relative strength. High Achievers showed comparatively high correct-response proportions across most tasks, whereas Structured Thinkers showed relatively stronger performance on tasks 7 and 8 and relatively weaker performance on task 5. These profile descriptions reflect the same task-response variables used to construct the clusters. They should therefore be interpreted as descriptive features of the selected k-means solution, rather than as independent inferential evidence that the clusters differ on those tasks.

In addition, because each task involved multiple CT components, the profile labels should not be understood as measures of isolated CT subskills or as fixed learner categories.

For teachers, these findings suggest that CT assessment results can be used not only when designing or selecting CT-related tasks but also when planning learning activities for students with different performance patterns. The identified clusters indicate that students may benefit from different forms and levels of instructional support. Profile information can help teachers decide when additional modelling, scaffolding, peer discussion, feedback, or enrichment activities may be appropriate when students engage with tasks involving multiple CT components. In this way, the patterns may serve as a useful tool for differentiating instruction and supporting students' development in computational thinking.

However, the results of the test showed that the precise separation of subskills to predict the success of completing the CT test is not possible; it confirms that subskills do not exist in isolation but form a mutually supportive system. The same findings have also become clear in another research (Dagienė et al., 2017; Román-González et al., 2018). Similarly to the study conducted by Tsai and others (2022), it can be said that while the levels of all subskills are dependent on each other, abstraction and decomposition are the key elements of predicting the success of solving the CT tasks. At the same time, as decomposition is reported as the most difficult (Selby, 2015), it can be said that the overall success in CT skills requires a complex and holistic approach. Also, it is possible to agree with the finding that students develop CT capabilities through multiple pathways, with varied patterns leading to different problem-solving approaches (Guggemos et al., 2023). For curriculum designers, the findings highlight the interconnected nature of CT components in authentic problem-solving tasks. Both the task analysis and the identified learner patterns suggest that students often engage with multiple CT components simultaneously rather than applying them in isolation. This indicates the value of curriculum materials that provide opportunities to work with different combinations of CT components across a variety of contexts and task formats.

Regarding cluster differences by grade and gender, our analysis indicated a significant association between grade and cluster membership: the High Achievers cluster included a larger share of 9th-grade students, whereas other clusters drew more evenly from 7th–8th grades. The comparison by gender showed that girls were more frequently represented in the High Achievers cluster, while boys were comparatively more present in the Low Achievers cluster. Pairwise analysis showed no significant differences between Pattern Recognisers and Low Achievers, nor between Low Achievers and Structured Thinkers. However, these associations should be interpreted cautiously rather than as stable grade- or gender-pattern findings.

When investigating the third research question (“Is there a statistically significant relationship between students' belonging to a specific cluster and their expressed intention to pursue further studies in the IT field?”), statistically significant between-cluster differences emerged in the two self-reported intention ratings. High Achievers reported statistically significantly higher mean ratings for both intention to continue IT studies and intention

to pursue IT-related careers compared to all other clusters. This finding provides empirical support for the claim that stronger CT task-response patterns could be associated with stronger self-reported intentions to continue in IT. However, the corresponding effect sizes were small ($\eta^2 = .033$ for intention to pursue IT-related studies and $\eta^2 = .044$ for intention to pursue a future in IT), suggesting that cluster membership was related to IT intentions, but the magnitude of this relationship remained modest. Moreover, these indicators do not show actual later study choices or career trajectories, and the study design does not allow causal claims. Prior work suggests that such intentions are shaped by multiple motivational components, including interest, self-efficacy, outcome expectations, and perceived task value/cost (Eccles & Wigfield, 2020; Zheng et al., 2024). Accordingly, our findings should be viewed as an initial signal that high CT task-response patterns co-occur with stronger self-reported IT intentions, rather than as a definitive account of the mechanisms linking CT to students' future study or career choices.

While direct empirical studies linking CT skills to IT career intentions remain limited, this finding aligns with broader theoretical frameworks in educational psychology and career development. Self-efficacy theory (Bandura, 1997) claims that individuals' beliefs about their capabilities to successfully perform tasks in a particular area significantly influence their career-related choices and persistence. Students who have developed stronger CT task-response patterns (for example, the High Achievers cluster in our research) might have a higher sense of self-efficacy for IT-related tasks, thereby increasing their willingness to envision IT as a viable career pathway. However, this study did not measure self-efficacy directly and did not follow students into later educational choices. Therefore, the present results should be interpreted as an association between task-response patterns and self-reported intentions, not as evidence that CT performance by itself determines students' choices to pursue IT studies and careers. The observed association nevertheless aligns with previous research showing that students with greater perceived abilities in mathematics and computing are more likely to express intentions to pursue STEM careers (Angeli & Giannakos, 2020; Chen et al., 2021; Zheng et al., 2024; Wang et al., 2024).

Importantly, no significant differences in IT career intentions were observed among the other three clusters (Pattern Recognisers, Low Achievers, and Structured Thinkers). This finding might have an important implication for educational practice. If the goal is to broaden the pipeline of students interested in IT careers, it may not be sufficient to focus solely on introducing CT concepts. Students may also benefit from learning experiences that support successful engagement with CT tasks and help them recognize their strengths and progress. Formative feedback, differentiated support, and encouragement may contribute to this process by helping students build confidence and connect their current learning to realistic next steps in informatics. Teachers and school-level informatics coordinators therefore have an important role in ensuring that students outside the High Achievers cluster receive support and guidance instead of signals that IT pathways are not for them.

The study has some limitations that need to be emphasized when generalizing the findings. First, the sample (all students from 30 Estonian schools) limits generalizability beyond the Estonian context, as CT skills might depend on the educational system and curriculum.

In addition, students were nested within 30 schools, but the analyses did not account for school-level clustering. Unmeasured school characteristics, including differences in the provision or content of informatics and CT-related instruction, may have contributed to students' CT performance patterns and to the observed associations with gender, grade level, and IT-related intentions. Future research should collect relevant school-level measures and use multilevel analytical approaches. However, as the participating schools form about 10% of Estonian schools, it gives an accurate country-specific overview which can be considered when comparing different approaches to CT skills in different education systems. In addition, the group from 9th grade was comparatively small and gender imbalanced. Also, the gender-related findings may reflect unmeasured factors. When evaluating students' CT skills, using several methods (students' description about the steps they take to reach a solution, for example) could help to make sure the results are in accordance with students' competence, as pure testing might not give a comprehensive view of that. Another limitation is that both intention variables were assessed with a single self-report item. While this was sufficient for an initial exploratory analysis, such measures may not fully capture the complexity of IT-related study and career intentions. Future research should build on this starting point by using more comprehensive multi-item measures.

6. Conclusion

Overall, this study contributes to understanding how lower-secondary students' computational thinking performance can be described not only through total test scores but also through meaningful patterns of task responses. It also highlights the potential relevance of CT competence for students' future educational orientation toward the IT field, while emphasizing that such relationships should be interpreted as correlational.

The study shows that Estonian lower-secondary students' performance on Bebras-based CT tasks can be meaningfully characterized through distinct task-response patterns. Rather than reflecting isolated CT subskills, these clusters captured patterns of performance across tasks that involved multiple CT components. Four clusters were identified, and students in the High Achievers cluster reported stronger intentions to pursue future studies and careers in IT. This finding suggests that students' CT task-response patterns and students' educational aspirations in IT are related, although the present study does not allow conclusions about causality.

The findings have practical value for informatics education. For teachers, the clusters can support more targeted formative feedback by showing which types of CT tasks students manage well and where they may need additional support. The clusters may also help teachers design classroom activities and support strategies that are responsive to students' different patterns of performance. The findings also support the integration of CT tasks across different subjects, especially in schools where separate informatics lessons are limited or unavailable.

The study also has several limitations. The sample was limited to one national context, and the 9th-grade subgroup was less balanced in terms of gender composition. In addition,

students' IT-study intentions were measured through narrow self-report indicators, which do not show actual later study choices or career trajectories. Finally, the cluster solution was exploratory and descriptive, and its stability was not evaluated through resampling or replication in an independent sample. The four clusters should therefore be interpreted cautiously as sample-specific task-response patterns.

Future research should examine CT task-response patterns in larger and more balanced samples, include longitudinal data to investigate whether such patterns predict later IT-related study choices, and explore how classroom interventions can support students in different CT patterns. Further studies could also combine task-response data with qualitative evidence from students and teachers to better understand how students approach multi-component CT tasks and how teachers can use this information for instruction.

Declarations

Author contributions. Conceptualization: RS, PL, TP; Methodology: RS, PL, TP; Software: RS; Validation: PL, TP; Formal analysis: RS, PL, TP; Investigation: RS, PL, TP; Data curation: RS; Writing—original draft: RS; Writing—Preview & editing: PL, TP; Visualization: RS; Supervision: PL, TP.

Ethical Approval. Ethics body and approval (unblinded for editor):

- Ethics committee/body: Research Ethics Committee of the University of Tartu
- Approval ID: 385/T-22
- Approval date: 2023-12-18.

Use of Artificial Intelligence Tools. The authors used generative AI tools in the final phase of the writing for language correction. No generative AI tools were used in other phases, and the authors take full responsibility for the manuscript.

References

- Adeoye, M. A., & Prastikawati, E. F. (2024). Simplifying the puzzle: How computational thinking and abstraction can help teachers conquer classroom complexity. *International Journal of Research in Education*, 4, 140–148.
- Angeli, C., & Giannakos, M. (2020). Computational thinking education: Issues and challenges. *Computers in Human Behavior*, 105, 106185.
- Ardito, G., Czerkawski, B., & Scollins, L. (2020). Learning computational thinking together: Effects of gender differences in a collaborative middle school robotics program. *TechTrends*, 64, 373–387.
- 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.
- Atman Uslu, N. (2023). How do computational thinking self-efficacy and performance differ according to secondary school students' profiles? The role of computational identity, academic resilience, and gender. *Education and Information Technologies*, 28(12), 6115–6139. <https://doi.org/10.1007/s10639-022-11425-6>
- Australian Curriculum, Assessment and Reporting Authority. (2024). Australian curriculum: Digital technologies. <https://www.australiancurriculum.edu.au/f-10-curriculum/technologies/digital-technologies/>
- Aydeniz, M. (2018). Integrating computational thinking in school curriculum. In M. S. Khine (Ed.), *Computational thinking in the STEM disciplines*. Springer.
- Balanskat, A., Engelhardt, K., & Ferrari, A. (2017). The integration of computational thinking (CT) across school curricula in Europe. *European Schoolnet Perspective*, 2. European Schoolnet.
- Bandura, A. (1997). Self-efficacy: The exercise of control. W. H. Freeman.
- Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K–12: What is involved and what is the role of the computer science education community? *ACM Inroads*, 2(1), 48–54. <https://doi.org/10.1145/1929887.192990>
- Bebras. (n.d.). Bebras international challenge on informatics and computational thinking. Retrieved August 22, 2026, from <https://www.bebbras.org/>
- Bocconi, S., Chiocciariello, A., Kampylis, P., Dagienė, V., Wastiau, P., Engelhardt, K., Earp, J., Horvath, M. A., Jasutė, E., Malagoli, C., Masiulionytė-Dagienė, V., & Stupurienė, G. (2022). Reviewing computational thinking in compulsory education. Publications Office of the European Union.
- Bounou, A., Lavidas, K., Komis, V., Papadakis, S., & Manoli, P. (2023). Correlation between high school students' computational thinking and their performance in STEM and language courses. *Education Sciences*, 13(11), 110.
- Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. *Proceedings of the Annual Meeting of the American Educational Research Association*.
- Cachero, C., Barra, P., Santiago, M., & López, G. (2020). Impact of programming exposure on the development of computational thinking capabilities: An empirical study. *IEEE Access*, 8.
- Carretero Gómez, S., Vuorikari, R., & Punie, Y. (2017). DigComp 2.1: The digital competence framework for citizens with eight proficiency levels and examples of use. Publications Office of the European Union. <https://doi.org/10.2760/38842>

- Chan, S., Looi, C., & Sumitono, B. (2020). Assessing computational thinking abilities among Singapore secondary students: A Rasch model measurement analysis. *Journal of Computers in Education*, 8, 213–236.
- Chen, C., Sonnert, G., Sadler, P., & Malan, D. (2021). Computational thinking and assignment resubmission predict persistence in a computer science MOOC. *Journal of Computer Assisted Learning*, 36(5), 581–594.
- Chongo, S., Osman, K., & Nayan, N. A. (2020). Level of computational thinking skills among secondary science students: Variation across gender and mathematics achievement. *Science Education International*, 31(2), 159–163.
- Computer Science Teachers Association. (2017). CSTA K–12 computer science standards. <https://www.csteachers.org/page/standards>
- Corrales-Álvarez, M., Ocampo, L. M., & Cardona Torres, S. A. (2024). Instruments for evaluating computational thinking: A systematic review. *Tecnológicas*, 27(59), e2950.
- Cui, Z., Ng, O., & Jong, M. S.-Y. (2023). Integration of computational thinking with mathematical problem-based learning: Insights on affordances for learning. *Educational Technology and Society*, 26(2), 131–146.
- 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.
- Dagienė, V., & Sentance, S. (2016). It's computational thinking! Bebras tasks in the curriculum. *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives*, 28–39.
- Dagienė, V., Sentance, S., & Stupurienė, G. (2017). Developing a two-dimensional categorization system for educational tasks in informatics. *Informatica*, 28, 23–44.
- Department for Education. (2014). National curriculum in England: Computing programmes of study. <https://www.gov.uk/government/publications/national-curriculum-in-england-computing-programmes-of-study>
- Didem, A. S. (2020). Investigating computational thinking skills based on different variables and determining the predictor variables. *Participatory Educational Research*, 7, 102–114.
- Dohn, N. B., Kafai, Y., Mørch, A., & Ragni, M. (2022). Survey: Artificial intelligence, computational thinking and learning. *KI - Künstliche Intelligenz*, 36(1), 5–16.
- Durak, H. Y., Karaoglan Yilmaz, F. G., & Yilmaz, R. (2017). Examining the relationship between digital game preferences and computational thinking skills. *Contemporary Educational Technology*, 8(3), 359–369.
- Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61, 101859. <https://doi.org/10.1016/j.cedpsych.2020.101859>
- Edelsbrunner, P. A., Simonsmeier, B. A., & Schneider, M. (2025). Cronbach's alpha of domain-specific knowledge tests before and after learning: A meta-analysis of published studies. *Educational Psychology Review*, 37, Article 4. <https://doi.org/10.1007/s10648-024-09982-y>
- European Commission. (2020). Digital education action plan (2021–2027): Resetting education and training for the digital age. European Union. europa.eu
- European Commission. (2021). Europe's digital decade: Digital targets for 2030. European Commission.
- Finnish National Board of Education. (2014). National core curriculum for primary and lower secondary (basic) education.

- Google Research. (n.d.). Exploring computational thinking. Google. <https://edu.google.com/resources/programs/exploring-computational-thinking/>
- Grover, S. (2017). Assessing algorithmic and computational thinking in K–12: Lessons from a middle school classroom. In P. J. Rich & C. B. Hodges (Eds.), *Emerging research, practice, and policy of computational thinking*. Springer.
- Grover, S. (2021). Toward a framework for formative assessment of conceptual learning in K–12 computer science classrooms. *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE '21)*, 31–37. <https://doi.org/10.1145/3408877.3432460>
- Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. *Educational Researcher*, 42(1), 38–43.
- Guggemos, J. (2021). On the predictors of computational thinking and its growth at the high-school level. *Computers and Education*, 161.
- Guggemos, J., Seufert, S., & Román-González, M. (2023). Computational thinking assessment: Towards more vivid interpretations. *Technology, Knowledge and Learning*, 28(2), 539–568. <https://doi.org/10.1007/s10758-022-09649-4>
- Haridussilm. (2023). Haridusandmete portaal [Education data portal]. Ministry of Education and Research. <https://haridussilm.ee/>
- Hava, K., & Koyunlu Ünlü, Z. (2021). Investigation of the relationship between middle school students' computational thinking skills and their STEM career interest and attitudes toward inquiry. *Journal of Science Education and Technology*, 30, 484–495.
- International Society for Technology in Education. (2019). *ISTE standards: Computational thinking competencies*. iste.org
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- Katai, Z., Osztian, E., & Lorincz, B. (2021). Investigating the computational thinking ability of young school students across grade levels in two different types of Romanian educational institutions. *Journal of New Approaches in Educational Research*, 10(2).
- Kobras. (n.d.). Kobras – informaatikaviktoriin koolidele. Retrieved August 22, 2026, from <https://kobras.eio.ee/>
- Kocak, O., Coban, M., Aydin, A., & Cakmak, N. (2021). The mediating role of critical thinking and cooperativity in the 21st century skills of higher education students. *Thinking Skills and Creativity*, 42.
- Liu, Z., Gearty, Z., & Richard, E. (2024). Bringing computational thinking into classrooms: A systematic review on supporting teachers in integrating computational thinking into K–12 classrooms. *International Journal of STEM Education*, 11, 51.
- Luik, P., Lehesaar, V., & Saadjärv, R. (2023). Teaching informatics in lower-secondary schools: Views of informatics teachers [Conference presentation]. *IEEE Global Engineering Education Conference*.
- Ma, H., Wang, X., Jin, Y., & Sun, J. (2021). The impact of different types of scaffolding in project-based learning on girls' computational thinking skills and self-efficacy. *Proceedings of the Tenth International Conference of Educational Innovation through Technology (EITT 2021)*, 362–366. IEEE.
- Mannila, L., Grgurina, N., Dagienė, V., Mirolo, C., Demo, B., Rolandsson, L., & Settle, A. (2016). Computational thinking in K–9 education. *Communications of the ACM*, 59(4), 31–33.

- Mertens, A. J., & Colunga, E. (2025). Assessing cognitive components of computational thinking. *Frontiers in Psychology*, 16. <https://doi.org/10.3389/fpsyg.2025.1434453>
- Ministry of Education and Research. (2023). The national curriculum of Estonia (in Estonian).
- Nakić, J., Rogulj, I., & Maleš, L. (2025). Rethinking the Bebras Challenge in virtual reality: Implementation and usability study of a computational thinking game. *Multimodal Technologies and Interaction*, 9(6), 60.
- Nouri, J., Zhang, L., Mannila, L., & Noren, E. (2020). Development of computational thinking, digital competence and 21st century skills when learning programming in K–9. *Education Inquiry*, 11(1), 1–17. <https://doi.org/10.1080/20004508.2019.1627844>
- Palts, T. (2021). A model for assessing computational thinking skills [Doctoral dissertation, University of Tartu]. University of Tartu Digital Archive ADA. <https://hdl.handle.net/10062/71913>
- Palts, T., & Pedaste, M. (2020). A model for developing computational thinking skills. *Informatics in Education*, 19(1).
- Piedade, J., Dorotea, N., Pedro, A., & Matos, J. F. (2020). On teaching programming fundamentals and computational thinking with educational robotics: A didactic experience with pre-service teachers. *Education Sciences*, 10, 214.
- Román-González, M., Moreno-León, J., & Robles, G. (2018). Can computational talent be detected? Predictive validity of the computational thinking test. *International Journal of Child-Computer Interaction*, 18, 47–58. <https://doi.org/10.1016/j.ijcci.2018.06.004>
- 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.
- Selby, C. C. (2015). Relationships: Computational thinking, pedagogy of programming, and Bloom's taxonomy. *Proceedings of the Workshop in Primary and Secondary Computing Education*. <https://doi.org/10.1145/2818314.2818315>
- Selby, C., & Woollard, J. (2013). Computational thinking: The developing definition. University of Southampton.
- Shute, V., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, 22, 142–158.
- Şişman, B., Hopcan, S., Polat, E., & Kucuk, S. (2021). A comprehensive assessment of secondary school students' computational thinking skills. *British Journal of Educational Technology*, 52, 1965–1980.
- Sousa, D. A., & Tomlinson, C. A. (2011). *Differentiation and the brain: How neuroscience supports the learner-friendly classroom*. Solution Tree Press.
- Stella, M., Kapuza, A., Cramer, C., & Uzzo, S. (2021). Mapping computational thinking mindsets between educational levels with cognitive network science. *Journal of Complex Networks*, 9(6).
- Strawhacker, A., & Bers, M. U. (2019). What they learn when they learn coding: Investigating cognitive domains and computer programming knowledge in young children. *Educational Technology Research and Development*, 67(3), 541–575. <https://doi.org/10.1007/s11423-018-9622-x>
- Sun, L., Hu, L., & Zhou, D. (2022). Programming attitudes predict computational thinking: Analysis of differences in gender and programming experience. *Computers and Education*, 181.
- Sun, L., Hu, L., Zhou, D., & Yang, W. (2023). Evaluation and developmental suggestions on undergraduates' computational thinking: A theoretical framework guided by Marzano's new taxonomy. *Interactive Learning Environments*, 31(10), 6588–6610.

- Sun, L., & Liu, J. (2023). Effects of gamified Python programming on primary school students' computational thinking skills: A differential analysis of gender. *Journal of Educational Computing Research*, 62, 846–874.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48, 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tariq, R., Aponte Babines, B. M., Ramirez, J., Alvarez-Icaza, I., & Naseer, F. (2025). Computational thinking in STEM education: Current state-of-the-art and future research directions. *Frontiers in Computer Science*, 6, 1480404.
- Teng, K., & Chung, G. K. W. K. (2025). Measuring children's computational thinking and problem-solving in a block-based programming game. *Education Sciences*, 15(1), 51. <https://doi.org/10.3390/educsci15010051>
- Tsai, M.-J., Liang, J.-C., Lee, S., & Hsu, C.-Y. (2022). Structural validation for the developmental model of computational thinking. *Journal of Educational Computing Research*, 60(1).
- Wang, C., Lu, C., Chen, F., Liu, X., Zhao, Q., & Wang, S. (2024). Growth mindset mediates the relationship between computational thinking and programming self-efficacy. *Education and Information Technologies*, 29, 21331–21354.
- Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., & Trouille, L. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology*, 25(1), 127–147.
- Werner, L., Denner, J., Campe, S., & Kawamoto, D. C. (2012). The fairy performance assessment: Measuring computational thinking in middle school. *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education (SIGCSE '12)*, 215–220. ACM.
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35. <https://doi.org/10.1145/1118178.1118215>
- Wing, J. M. (2011). Research notebook: Computational thinking—What and why? *The Link Magazine*. <https://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why>
- Wu, T.-T., Asmara, A., Huang, Y.-M., & Permata Hapsari, I. (2024). Identification of problem-solving techniques in computational thinking studies: Systematic literature review. *Sage Open*, 14(2). <https://doi.org/10.1177/21582440241249897>
- Yeni, S., Nijenhuis-Voogt, J., Saeli, M., Barendsen, E., & Hermans, F. (2024). Computational thinking integrated in school subjects: A cross-case analysis of students' experiences. *International Journal of Child-Computer Interaction*, 42, 100696.
- Zheng, J., Duffy, M., & Zhu, G. (2024). Predictors of university students' intentions to enroll in computer programming courses: A mixed-method investigation. *Discover Education*, 3, 144.