The Journal of Educational Research

ISSN: 0022-0671 (Print) 1940-0675 (Online) Journal homepage: www.tandfonline.com/journals/vjer20

Unplugged gamification in education: Developing computational thinking skills through embodied gameplay

Athanasios Christopoulos, Stylianos Mystakidis, Chrysostomos Stylios & Ioannis G. Tsoulos

To cite this article: Athanasios Christopoulos, Stylianos Mystakidis, Chrysostomos Stylios & Ioannis G. Tsoulos (18 Jun 2025): Unplugged gamification in education: Developing computational thinking skills through embodied gameplay, The Journal of Educational Research, DOI: 10.1080/00220671.2025.2517266

To link to this article: https://doi.org/10.1080/00220671.2025.2517266

Published online: 18 Jun 2025.

THE JOURNAL OF EDUCATIONAL RESEARCH
https://doi.org/10.1080/00220671.2025.2517266

Routledge
Taylor & Francis Group

Unplugged gamification in education: Developing computational thinking skills through embodied gameplay

Athanasios Christopoulos[a], Stylianos Mystakidis[b], Chrysostomos Stylios[c] and Ioannis G. Tsoulos[c]

[a]Turku Research Institute for Learning Analytics, University of Turku, Turku, Finland; [b]School of Natural Sciences, University of Patras, Patras, Greece; [c]Department of Informatics and Telecommunications, University of Ioannina, Arta, Greece

ABSTRACT

The study examines the effectiveness of unplugged gaming activities for stimulating Computational Thinking (CT). Departing from traditional, sedentary digital gamification, physically active alternatives were sought to reduce the potential health risks of excessive screen time. Additionally, efforts were made to engage girls. Using a mixed-methods approach, 311 adolescents (13–15 years old) were randomly assigned to an escape room or treasure hunt, where over a five-day summer vacation program, they participated in activities designed to improve CT components. Quantitative analysis revealed significant improvements in CT test scores after the intervention in both conditions for both genders. Qualitative analysis confirmed these findings and further highlighted their impact on engagement. Collaborative learning and cognitive challenge were positive aspects, while some challenges related to team dynamics and task complexity were also identified. The study demonstrates the potential of physically active gamification in CT education and highlights the importance of taking gender-specific experiences into account.

ARTICLE HISTORY
Received 5 April 2025
Revised 22 May 2025
Accepted 4 June 2025

KEYWORDS
Computational thinking; escape room; gamification; gender differences; treasure hunt

Introduction

Education is seen as crucial to preparing individuals, ensuring they have the knowledge and skills that contribute to society and the economy (Vonitsanos et al., 2024). Consequently, the learning objectives and educational methods need to be adapted to the developments that take place in the world. In the era we currently live in, digital technology and Artificial Intelligence are affecting our lives more than before, resulting in a society where we are interconnected through the Internet, have a rapid flow of information, and have the ability to use smart, connected devices to facilitate our daily lives as well as to foster further innovation (Vonitsanos et al., 2024). This has made computational thinking a vital ability that is being incorporated into educational programs worldwide (Belmar, 2022).

Computational Thinking (CT) concerns the ability to develop algorithmic solutions to complex problems (Ezeamuzie & Leung, 2022). Essential components of this ability are *abstraction*, *decomposition*, *pattern recognition*, and *algorithmic thinking*. The usual approach to teaching individuals CT concepts is through programming, including drag-and-drop interfaces that familiarize them with visual programming languages and later introduce them to syntax and writing codes for text-based programming (the latter usually in secondary education) (Sun et al., 2024). Computational education is usually "hands-on" and game-based, allowing students to learn through error and emphasizing the value of being able to debug (Bocconi et al., 2022). However, such educational methods contrast with the desire to unplug, disconnect adolescents from the Internet and digital devices as a means of reducing the physical and mental health risks of over-use which include myopia, obesity, and psychological distress (Pardhan et al., 2022; Rocka et al., 2022; Santos et al., 2023).

The present study, therefore, is focused on alternative, unplugged gamified ways to foster CT in adolescents through physical activities, such as escape rooms or treasure hunts.

Theoretical framework

When elements of games are used to help students engage with educational content, this is called gamification (Krath et al., 2021). From motivation theories, gamification is thought to decrease negative affect associated with more formal ways of learning (e.g., frustration, boredom) and increase engagement and motivation (Krath et al., 2021; Triantafyllou et al., 2024a). According to Self-Determination Theory, gamification satisfies the basic psychological needs that foster intrinsic motivation in the following ways: *(a)* it contributes to felt autonomy as it provides students control over the process and actions, *(b)* it increases feelings of competence by providing clear objectives and results, and *(c)* it can increase feelings of relatedness, although this highly depends on the social context provided (e.g., teamwork, or individual activities) (Luarn et al., 2023). Behavioral theories further

CONTACT Athanasios Christopoulos atchri@utu.fi Turku Research Institute for Learning Analytics, University of Turku, Turku, Finland.

explain students' active approach, feeling stimulated by extrinsic reinforcement, such as rewards and leaderboards (Hellín et al., 2023). Concerning learning, gamification is thought to have an advantage if it provides a context wherein students can learn by discussing their ideas with others or learning from others, including by observation (Krath et al., 2021; Triantafyllou et al., 2024a). Moreover, it often allows direct feedback, which can be key to facilitating learning (Krath et al., 2021; Triantafyllou et al., 2024a). This aligns also with constructivist learning theories affirming that high quality educational experiences leading to deeper comprehension and longer retention contain cognitive challenges and meaningful social interactions resulting in a positive emotional climate (Mystakidis et al., 2021).

Gamification seems particularly important for CT. A study among 7th graders showed that students do not always have positive attitudes toward programming. Especially among girls, these attitudes were more negative, even though their average CT skills were better than boys on average (Sun et al., 2022). Programming attitudes, in turn, were found to predict the development of CT (Sun et al., 2022). The more negative female attitudes toward CT are also found in other studies and align with the more general lower interest for Science, Technology, Engineering, And Mathematics (STEM) among females (Torres-Torres et al., 2024).

In the context of CT, gamified learning is almost automatically assumed to be computer- or otherwise digital-device-based. For example, a well-known international initiative is the Bebras competition, which provides interactive online challenges (e.g., puzzles, logic problems) that stimulate decomposition, pattern recognition, abstraction, and algorithm design learning among students of different ages (Triantafyllou et al., 2024b). Nevertheless, several studies have demonstrated the positive learning effects of unplugged gamified activities on CT. For example, a study among 7th-grade students used a board game in which the students needed skills necessary for CT, including planning their play figure's moves by decomposing a larger construction task into smaller steps and using instruction cards in sequences to create a path and collect resources (i.e., creating algorithms) (Kuo & Hsu, 2020). The aforementioned board game fostered the students' CT, especially when they worked on precise, pre-assigned tasks (Kuo & Hsu, 2020). A meta-analysis of unplugged activities, based on 13 studies, revealed a large effect on fostering CT in K-12 education (Chen et al., 2023). Nonetheless, these and other studies that Chen et al. (2023) included in their more extensive systematic literature review were most often investigating the effect of board games or unplugged computing challenges and competitions, such as Bebras.

Present study

While previous research has confirmed that unplugged gamified activities can foster students' CT, there are two critical research gaps. First, prior studies have focused on largely sedentary activities (Chen et al., 2023). While these activities may offer the benefit of not further increasing students' time on digital devices, other ways of gamifying learning activities are needed to stimulate more physical activity. Second, it is still unknown if gender affects the results obtained when using gamified activities to foster CT (Chen et al., 2023).

The aim of the present study is to investigate the potential of active, unplugged, gamified activities for secondary school students' CT (boys and girls). Two specific approaches were selected based on the literature: escape room and treasure hunt (Fotaris & Mastoras, 2019; Henriksen et al., 2021; Martia & Nurafni, 2023; Pazakou et al., 2025; Snelson, 2022). Although previously used for a variety of other educational objectives, both approaches seem to be experienced by students as enjoyable and immersive (Fotaris & Mastoras, 2019; Snelson, 2022). These approaches also stimulate teamwork, creativity, and critical thinking (Fotaris & Mastoras, 2019; Henriksen et al., 2021). Important for the present study was that they can be presented in an unplugged, physically active, manner. It was, thus, hypothesized that both approaches would result in increased CT knowledge acquisition, following the evidence found for other unplugged approaches (Chen et al., 2023). Furthermore, following the findings of previous studies using gamification approaches for educational purposes, no gender differences were expected in the learning outcomes (Veldkamp et al., 2021; Zahedi et al., 2021). As a secondary aim, competence beliefs and experiences were also explored. After all, competence beliefs and positive feelings can positively impact motivation (Ryan & Deci, 2000). Finally, given the limited evidence base, these feelings, and especially the in-depth information about the experiences, can be informative for future researchers and curriculum designers who desire to select and fine-tune the best activities to teach students CT.

The research questions guiding this study were:

RQ1: Do gamified unplugged activities (treasure hunt and escape room) significantly improve college students' computational thinking (CT) knowledge?

RQ2: How do gender and the type of gamified activity (treasure hunt *vs.* escape room) influence computational thinking improvement and students' affective experiences (competence beliefs, tension, negative affect)?

RQ3: What are students' and game masters' qualitative experiences regarding the computational thinking challenges and the gamification elements of the activities?

Materials and methods

Participants

Participants were 311 secondary school students aged 13–15 years old. There were 156 students in the treasure hunt condition and 155 students in the control condition. In each group, about half of the students were boys (48.7% in the treasure hunt and 47.1% in the escape room). The age distribution was the same in both groups (mean = 14.80, standard deviation = 0.80).

Procedure

The study took place at a summer camp, run by the local Orthodox Christian parish, in Greece. The camp environment prohibited the use of digital devices, necessitating the use of non-digital tools for all activities. Participants were divided into four cohorts, each scheduled for a one-week stay at the camp. Some participants departed earlier than planned due to family commitments or illness, and their data were not used in the study. Within each cohort, participants were further divided into subgroups based on their room (cabin) allocations. Participants were randomly assigned to one of the two gamified activity conditions: escape room or treasure hunt.

The educational program spanned five days and was structured in two distinct phases: *(a)* the instructional phase and *(b)* the applied phase (Figure 1). The instructional phase took two days. The first day, the researcher-educator provided an overview of the program placing particular emphasis on the element of collaboration and exploration. Additionally, participants were encouraged to ask questions and express their expectations. Thereafter, the assessment was administered to measure baseline CT skills. After the baseline assessment structured lessons were delivered on core CT concepts, including definitions, real-world examples, and interactive discussions. On the second day, the participants engaged in hands-on activities to reinforce CT concepts. These were logic puzzles and brainteasers requiring decomposition, pattern recognition, and algorithmic thinking through teamwork-based activities involving building structures and applying CT principles.

The applied phase took place on the third and fourth days and consisted of gamified activities (3 h per session). In the escape room condition, participants worked collaboratively to solve a series of interconnected puzzles and challenges within a multi-stage escape room scenario that required them to apply their CT skills. In the treasure hunt condition, participants engaged in deciphering clues, following logical paths, and using CT strategies to locate "hidden treasures." Table 1 presents the activities for both experimental conditions. While there were some minor differences in the specific learning objectives, both conditions focused on collaboration and teamwork and used activities of progressive complexity. On the fifth and final day, a post-intervention measurement and valuation took place. After having completed the individual tests and survey, a guided group discussion encouraged participants to reflect on their experiences, share insights, identify challenges, and discuss specific instances where they applied CT skills.

Instructional design

The escape room and treasure hunt formats were selected based on the research recommendations made by Hamari et al. (2016) on gamification in education, which demonstrate that challenge-based learning environments can increase student motivation and content retention.

The instructional design was guided by Shute et al. (2017) framework, which organizes CT into six key facets: decomposition, abstraction, algorithms, debugging, iteration, and generalization. To ensure pedagogical effectiveness, we incorporated the design principles from Huang and Looi (2021) on CT in K-12 education. In practical terms, we aimed at: *(a)* representing ideas in multiple ways, *(b)* supporting incremental and iterative development, *(c)* fostering debugging and error identification skills, and *(d)* promoting both individual reasoning and collaborative problem-solving. Importantly, each activity incorporates elements of productive failure (Kapur, 2008), where students encounter

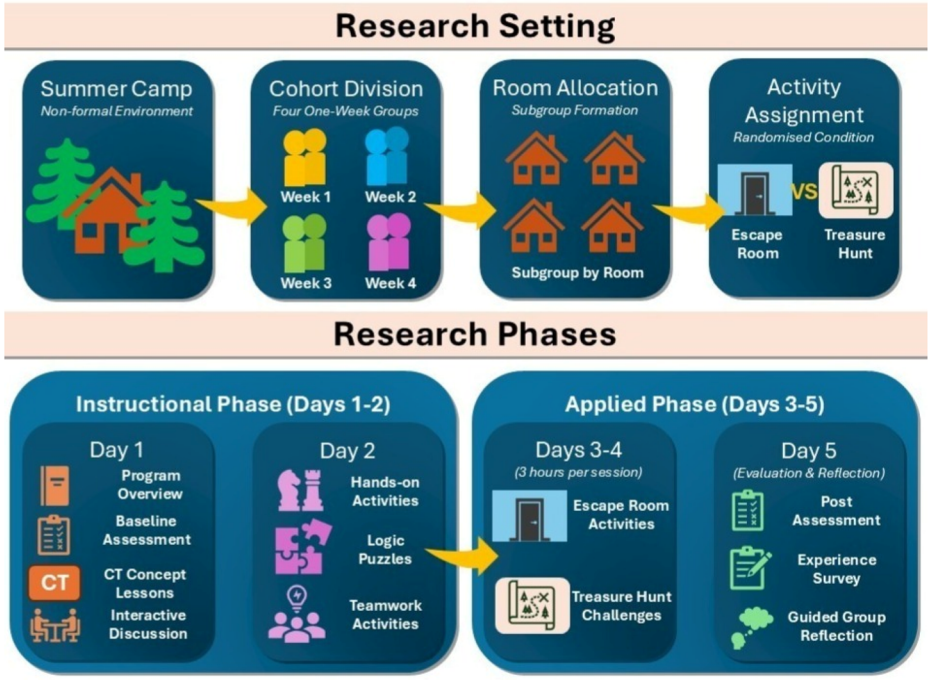

Figure 1. Graphic overview of the research design.

Table 1. Description of the activities.

Activity	Learning objective	Description	Inst. design principles*	Completion Time (est.)	Time limit
Escape room					
1. Binary code challenge	Binary code understanding, pattern recognition, collaborative problem-solving	Teams decode five 8-bit binary numbers to spell the word "CLOCK," identifying the location of a hidden envelope containing two more binary numbers that form a four-digit code to unlock the exit. Teams can use decoder charts with a 5-min time penalty.	PAT, B1, B2	20′	30′
2. Pattern lock decryption	Pattern recognition, sequence analysis, abstraction	Teams solve six distinct visual pattern puzzles, selecting the correct continuation (A, B, C, or D). Each letter corresponds to a specific number ($A=2$, $B=5$, $C=7$, $D=8$), forming a six-digit code to unlock the final challenge.	PAT, ABS, B2, B3	25′	35′
3. Logic circuit puzzle	Boolean logic application, step-by-step planning, logical thinking	Teams analyze a circuit with "AND" and "OR" gates, calculating outputs for all input combinations. The completed truth table reveals a binary sequence that unlocks a padlock protecting the key. Incorrect calculations result in invalid codes.	DEC, ALG, B3, B4	25′	35′
4. Network flow challenge	Network topology understanding, system thinking, component relationships	Teams match network device pictures to correct locations on a network diagram. Each device card reveals a number on its back. The correct device arrangement reveals a path code between the ISP and Computer 2 that unlocks the prize.	DEC, ABS, B3, B4	25′	35′
5. Algorithm assembly	Programming sequence understanding, algorithmic thinking, debugging	Teams write instructions to guide a blindfolded teammate through a floor maze using direction cards (move forward, turn left/right). Encountering walls requires restarting, while penalty squares deduct time. Bomb squares force algorithm recreation.	ALG, DEB, ITR, B5, B6	30′	40′
6. System architecture challenge	System design, component integration, abstraction	Teams reconstruct a modular computer system model where components must be connected in a functional configuration, demonstrating understanding of system architecture and relationships between hardware elements.	DEC, ABS, GEN, B5, B6	30′	40′
Treasure hunt					
1. Binary trail hunt	Binary code application, sequential thinking, map navigation	Teams follow a treasure hunt across campus by decoding binary numbers into letters (e.g., L-I-B-R-A-R-Y). Each location contains an envelope with the next binary clue, ultimately leading to the hidden treasure at the decoded location.	PAT, B2, B3	25′	35′
2. Logic path hunt	Logic gate operations, problem-solving, spatial reasoning	Teams solve logic gate puzzles at different campus locations, with each solution corresponding to a letter value. The sequence of correct solutions spells "STADIUM," leading teams to the final treasure location.	ALG, PAT, B3, B4	30′	40′
3. Network trial challenge	Network topology, data flow analysis, critical path identification	Teams arrange network device cards on a diagram. When correctly placed, the numbers on the backs of cards along the path from Modem→Router→Switch→PC1 form a code that indicates the treasure's location on a chart.	DEC, ABS, PAT, B3, B4	25′	35′
4. Code path challenge	Pattern identification, algorithmic application, abstraction	Teams navigate through a series of encoded messages using progressive substitution ciphers. Each decoded pattern reveals coordinates to the next challenge location, with complexity increasing at each stage.	PAT, ABS, ITR, B3, B4	30′	40′
5. Algorithm trail quest	Algorithm development, precise instruction design, optimization	Teams create step-by-step navigational algorithms to guide a blindfolded teammate from a starting classroom to treasure in the laboratory, following a grid-based floor plan. Hitting obstacles requires algorithm revision with time penalties.	ALG, DEB, ITR, B5, B6	35′	45′
6. System path challenge	Algorithm execution, optimization, system-level thinking	Teams follow pseudocode instructions to physically navigate between locations, requiring precise interpretation of commands and efficient pathfinding to reach the final treasure location within time constraints.	ALG, ITR, GEN, B4, B5, B6	35′	45′

**CT facets*: DEC: decomposition; ABS: abstraction; ALG: algorithms; DEB: debugging; ITR: iteration; GEN: generalization; PAT: pattern recognition. *Bloom's taxonomy levels*: B1: remembering; B2: understanding; B3: applying; B4: analyzing; B5: evaluating; B6: creating.

controlled challenges that promote deeper conceptual understanding through problem-solving persistence.

The progression of difficulty across levels follows Bloom's revised taxonomy (Anderson & Krathwohl, 2001), beginning with remembering and understanding binary representations (Level 1), advancing through application and analysis of logical circuits (Levels 2–3), and culminating in evaluation and creation of algorithmic solutions (Levels 4–6).

Assessment of CT acquisition is embedded within the activities through performance metrics (completion time, error frequency) and process observations that align with the assessment approaches that Brennan and Resnick (2012) recommend. The proposed assessment model allows instructors to identify misconceptions in real-time and provide targeted scaffolding.

Table 1 presents the integration of these principles into the CT activities, detailing the specific learning objectives, the activity descriptions, and the time parameters for implementation, whereas Figure 2 illustrates the camp layout and indicative examples of the CT challenges.

Instruments

The Appendices provide all instruments used in this study.

Quantitative data collection

Computational Thinking was measured with three different tests. The tests were developed for the purpose of this study, inspired by several online resources[1] as well as formal educational resources dedicated to unplugged programming.[2] The first test consisted of 10 multiple-choice questions, each presenting the correct answer among three wrong alternatives. For example, the students were asked "What is a 'pseudo-code'?" with the potential answers: (a) It is a type of programming language; (b) It is a simple way of expressing a program or algorithm that uses a mixture of English phrases and syntax resembling programming languages (correct answer); (c) It is a form of shorthand notation for representing binary data; or (d) it is a type of code used for debugging purposes. The second test consisted of 10 sentences, and the students needed to fill in the blanks (e.g., the purpose of _____ in a program is to store data that can be used by the program [correct answer: variables]). The third test involved matching 10 CT concepts with their correct definitions.

Competence Beliefs were measured with the scale developed by IJsselsteijn et al. (2013). The scale consists of 5 items (e.g., "I felt skillful") on a Likert scale from 0 = *not at* all to 4 = *extremely*. This scale seems appropriate to evaluate gamification experiences in young samples (Lampropoulos et al., 2023). The scale showed acceptable reliability in the present study (Cronbach's $\alpha = .74$).

Emotional Feelings were investigated using the **tension** and **negative affect** scales of the same authors (IJsselsteijn et al., 2013). Tension was operationalized by 3 items (e.g. I felt frustrated) and negative affect by 4 items (e.g., It gave me a bad mood). The same 0–4 Likert scale was used. Sufficient internal consistency was found for both tension ($\alpha = .65$) and negative affect ($\alpha = .74$).

Qualitative data collection

The post-intervention discussions were conducted in a semi-structured format by the team leaders and the lead researcher at the conclusion of the activities. The conversations focused on two main topics: the CT activities and the experiences with the educational games. Team leaders were instructed to encourage participation from all students, based on predefined prompts, while allowing for natural conversation flow.

For the topic of CT, the following prompt was used:

> *"Think back to the escape room/treasure hunt activity. Describe a time when you had to break down a complex puzzle or a challenge into smaller, easier-to-solve steps. How did this approach help you solve the puzzle? Did you use any strategies that you think a computer programmer might also use?"*

For the topic educational games, the prompt was:

> *"What aspects of the escape room/treasure hunt experience did you find the most fun or engaging? Did the competitive elements, puzzles, teamwork, or the story of the game make it more enjoyable? If you could design your own escape room or treasure hunt game, what features would you include to make it even more fun and interesting?"*

Each team leader used his/her mobile phone as a voice recording device to capture the discussions for later analysis. The data were later transcribed to facilitate the analysis.

Analysis

The study employed a mixed-methods approach. The quantitative analysis was carried out using SPSS (v.30). Before conducting the Repeated Measures ANOVA, the assumption of normality for the dependent variables was assessed, including the normality of the difference scores between the pretest and post-test measures. Visual inspection of histograms and Q-Q plots confirmed normality, and no severe skewness or kurtosis was found (Kim, 2013). For the evaluation of overall progress in CT from pretest to post-test, paired sample *t*-tests were utilized. This approach was chosen to directly assess significant changes in mean scores within the same group of students over time. Repeated measures of Analysis of Variance (ANOVA) were used when evaluating if gender (boys *vs.* girls) and type of activities (escape room *vs.* treasure hunt), as between-group variables, affected the mean scores obtained on the different assessments. This test was chosen given that the same participants were measured at multiple time points (pretest and post-test), allowing for the examination of within-subject changes over time while controlling for individual differences (Sharpe & Cribbie, 2023). Analysis of Variance (ANOVA) was used to examine the quantitative evaluations (competence beliefs, tension, negative affect) which were collected post-intervention. This approach allowed for the comparison of mean scores across different conditions and genders, with Bonferroni adjusted pairwise comparisons applied to control for Type I error rate in multiple comparisons (Field, 2018).

The qualitative data were analyzed using thematic analysis (Braun & Clarke, 2021). After thorough data familiarization (the first step of thematic analysis), the most representative quotes relevant to the research questions were organized by

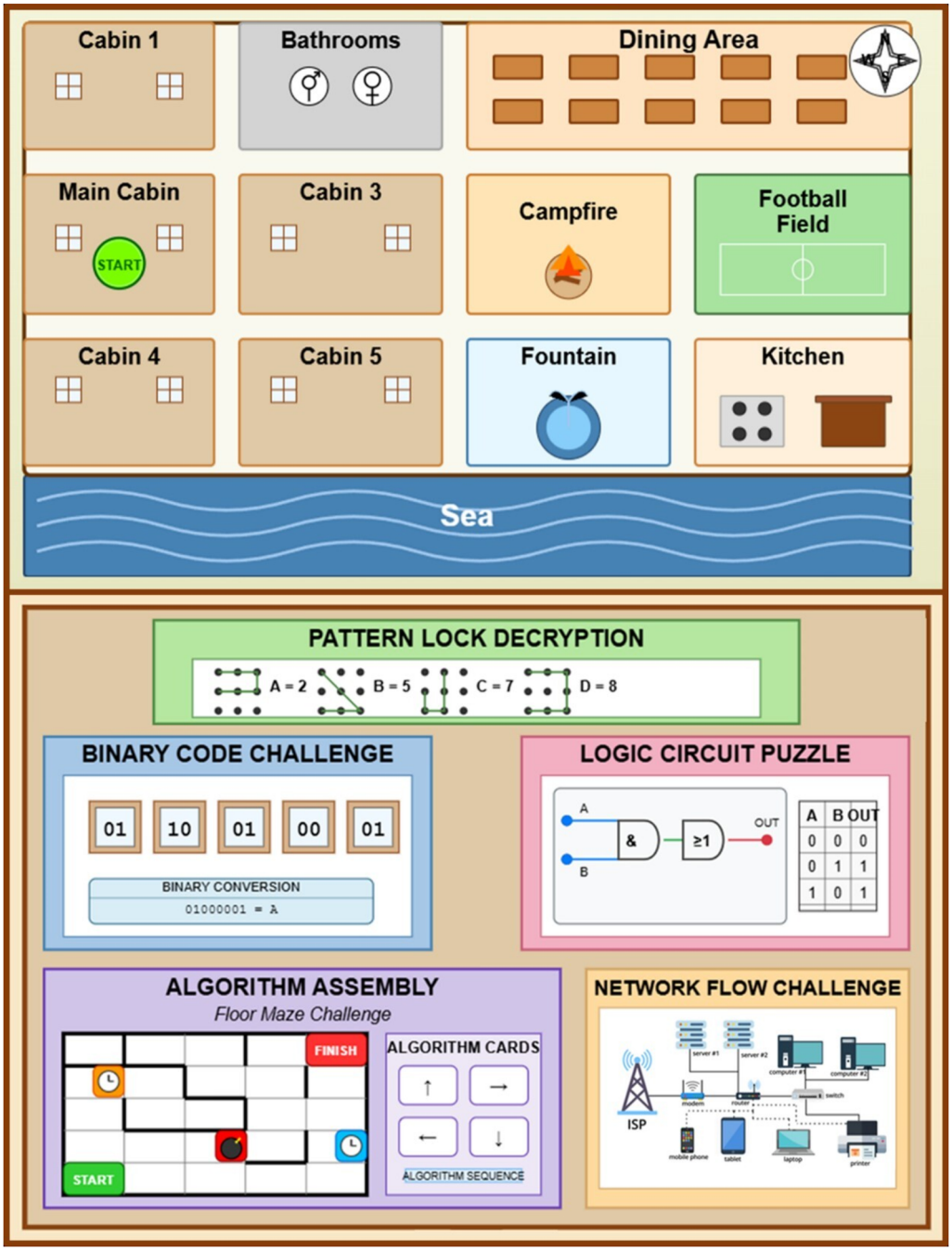

Figure 2. Overview of the summer camp layout (upper frame) and indicative visualizations of the CT activities.

topic and activity, as well as positive *vs.* negative for the students and game masters (i.e., team leaders) separately. To obtain objectivity, a second, independent researcher used these quotes (>100) to first derive codes (second step), labeling similar quotes when there were at least 3 that could be coded with the same label. Thereafter, from these codes, the larger themes were identified and labeled (third step). For Computational Thinking, themes included engaging and enjoyable activities, breaking down problems into smaller parts, and interest in further learning (positive feedback), alongside losing sight of the bigger picture and frustration (negative feedback). For Gamification, themes encompassed engagement and motivation, collaborative learning, and cognitive challenge (positive feedback), contrasted with progress-related stress, team dynamics issues, and unclear learning purpose and outcomes (negative feedback). The first researcher checked this analysis's relevance and correctness to ensure the themes formed an accurate description of the data (fourth step). There were no points of disagreement and the fifth and sixth steps (defining and naming themes and producing the report) were made using these themes.

Results

Effects on computational thinking

Table 2 shows the averages and standard deviations of the students on the pre-and post-tests of CT. On average,

students on the pretest knew more than 6 of the answers correctly, and on the multiple-choice test, they approached an average of 7. Paired sample *t*-tests confirmed that significant progress had been achieved at the end of the summer camp. On average, the students knew more than 7 of the answers correctly on the tests on the sentence completion and matching test, with a mean score of 7.50 on the multiple-choice test at the end of the summer camp. As shown in Table 2, the effect sizes for all three CT assessments were found to be medium to large (Multiple Choice: $d = .71$; Sentence Completion: $d = .65$; Matching: $d = .73$). On average, students improved by more than half to nearly three-quarters of a standard deviation in their CT knowledge.

Repeated measures of ANOVA on the multiple-choice questions revealed a large significant effect of time, $F(1, 307) = 159.36$, $p < .001$, $\eta_p^2 = .34$, confirming the students' computational knowledge improvement. There also was a small main effect of gender, $F(1, 307) = 12.36$, $p < .001$, $\eta_p^2 = .04$ which indicated that the CT scores of boys generally were better than those of girls. No other significant main or interaction effects were found, except for a three-way interaction of time, condition, and gender, $F(1, 307) = 4.65$, $p = .032$, $\eta_p^2 = .02$ (Figure 3). Boys and girls showed improvements in both conditions, although for boys the escape room ($d = .49$) seemed to be slightly more effective compared to the treasure hunt ($d = .69$). However, the difference between the treasure hunt and escape room groups of boys were insignificant at pretest as well as post-test, and the same was found for girls (although the pretest scores of girls in the treasure hunt condition were marginally lower, $p = .061$). The smaller effect of the treasure hunt for boys did lead to a significant overall difference in the post-scores between the escape room ($M = 7.68$, $SD = 1.64$) and the treasure hunt ($M = 7.31$, $SD = 1.57$), $t(309) = 2.07$, $p = .040$, $d = .23$. The difference between the scores of the escape room and treasure hunt groups had not been present at the pretest, with $M = 7.03$, $SD = 1.88$ and $M = 6.73$, $SD = 1.73$ respectively, $p = .142$.

Table 2. Mean pre- and post-test scores on computational knowledge.

Test	Pre-test *M* (*SD*)	Post-test *M* (*SD*)	*t*	Effect size *d*
Multiple choice	6.88 (1.81)	7.50 (1.61)	12.53***	.71
Sentence completion	6.27 (2.08)	7.12 (1.17)	11.41***	.65
Matching	6.34 (1.95)	7.01 (1.72)	12.82***	.73

***$p < .001$.

The same analysis was repeated for the sentence completion task. Only a main effect of Time was found, $F(1, 307) = 128.53$, $p < .001$, $\eta_p^2 = .30$. There were no other main or interactions effects of gender and/or condition. As such, it indicates that the students had higher scores on the sentence completion task at the end of the summer camp compared to before the summer camp, regardless of whether they had participated in the treasure hunt or escape room and regardless of their gender, with similar scores for boys and girls. On the final matching task, the same results were found. There only was an effect of Time, $F(1, 307) = 162.30$, $p < .001$, $\eta_p^2 = .35$.

Quantitative evaluations

Table 3 shows the means and standard deviations obtained on the quantitative evaluations. Within the potential range of scores from 0 to 4, the scores reflect that students, on average, felt competent and experienced little tension or negative affect. Analysis of Variance was used to examine the quantitative evaluations of the two conditions, with Bonferroni adjusted pairwise comparisons.

On **competence beliefs**, a main effect of Condition was found, $F(1,307) = 93.64$, $p < .001$, $\eta_p^2 = .23$, reflecting that students who had participated in the escape room had felt more competent. However, there were also significant effects for Gender ($F[1,307] = 318.73$, $p < .001$, $\eta_p^2 = .51$) and Gender x Condition ($F[1,307] = 90.54$, $p < .001$, $\eta_p^2 = .23$) with boys reporting more competence beliefs in the escape room. Interestingly, girls from both groups felt more competent compared to boys, $p < .001$.

Figure 3. Pre- and post-test scores on the multiple-choice test for boys and girls in the treasure hunt and escape room condition.

For **tension**, there was a main effect for Condition, $F(1,307) = 200.65$, $p < .001$, $\eta_p^2 = .40$, reflecting that students who had participated in the escape room had felt less tension. There were also significant effects for Gender ($F[1,307] = 6.12$, $p = .014$, $\eta_p^2 = .02$) and Gender x Condition ($F[1,307] = 18.86$, $p < .001$, $\eta_p^2 = .06$). Both boys and girls had felt less tense in the escape room compared to the treasure hunt. However, while there was no gender difference in the experienced tension during the treasure hunt ($p = .186$), boys had felt significantly more tense than girls in the escape room, $p < 001$.

Finally, on **negative affect** the strongest Condition effect was found, $F(1,307) = 446.28$, $p < .001$, $\eta_p^2 = .59$. Again, there also was a Gender($F[1,307] = 18.19$, $p < .001$, $\eta_p^2 = .06$) and Gender x Condition ($F[1,307] = 27.38$, $p < .001$, $\eta_p^2 = .08$) effect. Both boys and girls had felt more positively in the escape room compared to during the treasure hunt. However, girls in the escape room had felt more positively compared to boys, $p < .001$. In the treasure hunt, the feelings of boys and girls had been comparable.

Qualitative evaluations

The feedback given by the students in the discussion groups provided further insights into their experiences. Within the topic of CT, the positive feedback revealed three themes: *(1) engaging and enjoyable activities*, *(2) breaking down problems into smaller parts*, and *(3) interest in further learning*. The negative feedback was captured by two themes: *(1) losing sight of the bigger picture* and *(2) frustration*.

Positive feedback

Participants in both the *escape room* and the *treasure hunt* activities expressed overly positive views of the CT challenges. Many emphasized that they had found the activities **engaging and enjoyable**, with comments, such as "This was way more fun than regular school" (*escape room*) and "the activity was cool." Several participants appreciated the problem-solving aspects, comparing it to detective work (*escape room*) and enjoying the process of "figuring things out (*escape room* and *treasure hunt*)."

A common theme was the benefit of **breaking down problems into smaller parts**, which made complex tasks feel less difficult and more manageable. For example, one student in the *escape room* said: "Breaking down the puzzle into smaller parts helped a lot. It made it less scary and more fun." In the *treasure hunt* condition, participants similarly related the small steps to positive feelings, as reflected in their quotes: "Each small win made me feel more confident" and "Each step we solved felt like a team victory." This theme was also clear to the game masters. For example, one of them mentioned: "Students appreciated the clarity that a step-by-step approach provided, making it easier to track progress and solve puzzles systematically" (*escape room*).

While overall, the evaluations of the two conditions were highly similar, the participants in the *escape room* **showed interest in further learning**, with comments, such as "I want to learn more about programming now" and "can we do another escape room soon?" In the *treasure hunt*, this interest was not as explicit. This finding aligns with the somewhat more positive evaluations of the escape room compared to the treasure hunt.

Table 3. Mean scores and standard deviations on the quantitative evaluations of boys and girls.

Variable	Gender	Escape room	Treasure hunt	*p*
Competence beliefs	Boys	3.18 (0.39)	2.51 (0.34)	<.001
	Girls	3.47 (0.24)	3.47 (0.25)	.907
Tension	Boys	0.84 (0.47)	1.31 (0.51)	<.001
	Girls	0.52 (0.32)	1.40 (0.36)	<.001
Negative affect	Boys	0.86 (0.43)	1.49 (0.37)	<.001
	Girls	0.48 (0.24)	1.53 (0.34)	<.001

Note. *p*-Values are obtained using *post-hoc* Bonferroni-corrected pairwise comparisons.

Negative feedback

Despite students positively commenting on breaking down problems, in some cases, students experienced they were **losing sight of the bigger picture**, e.g., "I found it difficult to keep track of all the pieces we had solved. It got confusing quickly" (*treasure hunt*) and "I felt lost and didn't know where to start" (*escape room*). These comments indicate that the breaking down was sometimes confusing. A Game Master noted, "The iterative process of trial and error was sometimes frustrating for students who preferred more direct solutions" (*escape room*).

A second theme was **frustration**. In some cases, students had felt the tasks were too difficult. This was described in both conditions with words, such as "frustration" and "stressful." It seemed to decrease motivation to continue [e.g. "Some puzzles were too confusing and I gave up" (*escape room*)]. This theme clarifies that, while the activities seemed enjoyable for most students, some individual students may have felt it was above their competency or Zone for Proximal Development.

Effects of gamification

Within the topic of **gamification**, three themes captured positive feedback: *(1) engagement and motivation*, *(2) collaborative learning*, and *(3) cognitive challenge*. The negative feedback was presented by: *(1) progress-related stress*, *(2) team dynamics*, and *(3) unclear learning purpose and outcomes*.

Positive feedback

Overall, the adoption of gamification seemed successful in achieving **engagement and motivation**. The students expressed how they had immersed in the activities that they had experienced as enjoyable, as reflected by quotes like: "The story made it feel like a real adventure, not just school stuff" (*escape room*) and "Acting out the roles [team leader, team navigator, detective, scout] was hilarious and made everything more fun. It was like being in a real-life game" (*treasure hunt*). The activities transformed the learning experience into something desirable, "The whole thing felt like a game, not boring school work" (*escape room*). This theme was confirmed by the feedback provided by the Game

Masters, "The fun and interactive nature of the game made learning [about CT] enjoyable" (*escape room*). However, different from the student experience, some Game Masters had experienced stronger emergence in the treasure hunt, e.g., "The treasure hunt's focus on adventure and exploration provided a unique thrill and sense of discovery, different from the confined space of an escape room." On the other hand, the expression "The physical nature of the treasure hunt led to issues with pacing and fatigue, affecting students' ability to maintain focus and energy" provides some clue that this positive effect may not have been achieved to its full potential because of the physical challenges.

An element that received explicit attention, was that **collaborative learning** was evaluated positively. Several students in both conditions mentioned they had liked "working as a team." Some also referred to the team-competition element, e.g. "I loved trying to beat the other teams! It made it super exciting" (*escape room*). This was also emphasized by the Game Masters, e.g., "Many students appreciated the opportunity to work together, share ideas, and support each other" (*escape room*).

Finally, it became clear that the experience of **cognitive challenge** was evaluated positively. Highly similar quotes in both conditions, referred to thinking hard, such as: "The puzzles were really cool because we had to *think hard* and work together." (*escape room*) and "The puzzles were actually cool! We really *had to use our brains* and work as a team" (*treasure hunt*) [emphasis added]. The Game Master formulated this on the level of different approaches: "Students appreciated that some clues required sequential thinking, while others allowed for non-linear, creative approaches to find the treasure" (*treasure hunt*). It may be that the variety helped different students contribute, a possible reason or explanation as of why they may have emphasized the team element when providing positive feedback on it being cognitively challenging.

Negative feedback

Students who were not winning tokens or could not complete the tasks successfully reported **progress-related stress**. For example, one student said "Losing lives was stressful because we were scared to run out" (*escape room*). Another student who had participated in the *treasure hunt* explained "I felt stressed when we couldn't earn the tokens because some puzzles were just too hard." The competitive element may have contributed to these feelings. As explained by one student in the *treasure hunt*: "Seeing other teams ahead of us on the leaderboard made me feel kind of bad about our progress." The progress-related stress was also observed by the Game Masters "Some students felt that the competitive aspect created stress and sometimes overshadowed the learning objectives."

Further, **team dynamics** were not positively evaluated by all students, despite the fact that it was experienced positively in many cases. Complaints were mostly about some not contributing sufficiently, such as "Our team didn't work well together. Some people did all the work, others just watched" (*escape room*) and "Some people in our group didn't really participate. It felt like a few of us were doing all the work" (*treasure hunt*). Others were about not getting along, such as "I wish I could have chosen my team. We didn't really get along, so it was hard to work together." The Game Masters explained the potential underlying cause of mixed initial computational skills of the group members: "Variability in computational thinking skills among team members affected group performance and satisfaction" (*escape room*). In the *treasure hunt* condition, the physical space seemed to be yet another barrier: "Some students found it challenging to coordinate physically with team members, especially in large or complex spaces." Here, besides computational ability, physical ability also played a role: "Physical ability varied greatly, affecting how much each student could contribute to the hunt."

The theme **unclear purpose and learning outcomes** was labeled to reflect the fact that students did not always understand why they were participating. For example, a student expressed "The story was confusing. I wasn't sure what we were supposed to do or why." Some students also directly referred to not having learned anything, such as: "The escape room was okay, but I didn't really learn anything new." This was not how it was experienced by the Game Masters: "Many students recognized that using logic and structured thinking, similar to programming, helped them efficiently solve puzzles."

Discussion

The present study focused on the possibility of teaching students CT using physically active, unplugged activities with a mixed-method approach. The expectation that positive results would be obtained with treasure hunt and escape room activities among boys and girls was confirmed. Consistent and statistically significant improvement, coupled with these medium-to-large effect sizes, indicates that the gamified activities were indeed effective in tangibly enhancing students' CT knowledge. Furthermore, the qualitative analysis of group discussions revealed that both students and Game Masters were positive about the stimulation of CT and the use of gamification. Students experienced the activities as engaging and enjoyable, and especially found the problem decomposition valuable; breaking down problems into smaller parts made it more manageable and less intimidating. Nevertheless, some students lost their view of the bigger picture and/or felt frustrated when they found the tasks too difficult. The adoption of gamification practices appeared successful, as evidenced by students' and Game Masters' reports of engagement and motivation related to collaborative learning and cognitive challenges. On the other hand, progress-related stress, negative team dynamics, and unclear purpose or learning outcomes were mentioned less frequently.

While previous research had provided initial evidence concerning the use of mostly sedentary unplugged gamified activities to foster students' CT (Chen et al., 2023), the present study showed beneficial effects of treasure hunt and escape room activities, which were physically active forms

of gamification. The activities had been selected based on previous research showing their educational value and potential is stimulating student motivation (Fotaris & Mastoras, 2019; Henriksen et al., 2021; Snelson, 2022). Previous studies have not compared treasure hunt and escape room activities, however. The present study was able to provide this comparison by randomly assigning the students to either condition. The statistical analysis confirmed that the groups were statistically equivalent at baseline, which facilitated the interpretation of the observed patterns of improvement. While both activities can be used successfully for CT, a slight advantage to using escape rooms over treasure hunts was found. The escape room activities seemed slightly more effective for boys compared to the treasure hunt, and both boys and especially girls felt less tension and more positive about the overall experience. Moreover, from the qualitative findings, the desire to learn more about programming and participate in similar activities was also more clearly present among the students who had participated in the escape room activities. A possible explanation is that the physical space is more confined when using escape room activities, which can help students feel more connected and better protected by the oversight of the Game Masters. After all, according to Self-Determination Theory, relatedness helps students feel intrinsically motivated (Ryan & Deci, 2000). Moreover, the oversight may have resulted in the objectives and results being clearer during the escape room activities as the Game Masters were able to oversee the process quickly and were able to provide continuous communication. According to Self-Determination Theory, this can have fulfilled students' need for feeling competent better than in the treasure hunt, which in turn may have increased their motivation (Ryan & Deci, 2000).

Similar to previous studies using gamification approaches for educational purposes (Veldkamp et al., 2021; Zahedi et al., 2021), we found that both the treasure hunt and the escape room resulted in positive learning outcomes for boys and girls. Girls had even felt more competent in the escape room compared to boys. This finding has implications for CT education in terms of motivation. Research shows that girls tend to have less favorable attitudes toward CT compared to boys (Sun et al., 2022), aligning with less STEM interest among females (Torres-Torres et al., 2024). By using gamified approaches, girls may be motivated to learn CT and other STEM-related skills. Overall, CT in girls seemed to be similar to that of boys, except for a difference found in the multiple-choice exam before as well as after the intervention. This finding aligns with previous research showing that the male advantage in math and science is inflated when using multiple-choice questions (Griselda, 2024). According to the "paradox of choice," especially anxious or less confident students are likely to feel confused when presented with alternatives (Griselda, 2024). The fact that, in the fill-in-the-blank and the matching test, no gender differences were found suggests that girls may have underperformed on the multiple-choice exam because they felt insecure.

The themes of "losing sight of the bigger picture" and "frustration" point to a challenge in designing problem-solving tasks, especially those involving decomposition (breaking down problems into smaller parts). Although problem decomposition is a core CT skill (Ezeamuzie & Leung, 2022) and was welcomed by many students, the iterative process of solving smaller puzzles sometimes led to cognitive overload or a disconnection from the overarching problem. This suggests that the support provided may have been inadequate for some students: it left them struggling to reintegrate their solutions or maintain a coherent understanding of the larger challenge. The reported frustration suggests that the difficulty occasionally exceeded some students' optimal challenge level. According to theories of optimal challenge and flow, tasks that are too difficult can lead to anxiety and withdrawal rather than sustained engagement (Krath et al., 2021). A possible solution may be to include more explicit mechanisms for metacognitive reflection (e.g., regular check-ins or visual progress charts that help students contextualize their current task within the broader problem).

Further, while not as pervasive as the positive feedback, the presence of some negativity about the gamification elements indicates areas where the gamified design or facilitation could be refined. Progress-related stress suggests the competitive elements may have inadvertently created some anxiety for a few students. Issues with team dynamics highlight the need for more structured collaboration support, and the finding that some experienced unclear learning purpose and outcomes underscores the importance of explicitly bridging the gap between engaging game mechanics and the underlying educational objectives to ensure meaningful learning for all students.

Limitations and future work directions

The study naturally comes with its limitations. Although multiple CT tests were used, they focused more on knowledge than skill. The students showed improved skills during the activities, but as these were team efforts, no information about individual students' gains was obtained. From the qualitative data, there seemed to be a few students who found the activities difficult. Another limitation was that the intervention included an instructional and an applied phase, but only the effects of both together were assessed. It therefore remains uncertain which specific instructional and student activities were most helpful in obtaining the results. There also was no control group of students who were engaged in CT activities on digital devices. As such, it remains uncertain if the improvements made were similar or different from those that can be obtained with more standard CT lessons, such as programming, and writing code for text-based programming on the computer (Sun et al., 2024). Future studies may wish to include a control group and add measures to more completely assess CT skills in practice on the individual level. In future studies, it may also be assessed if activities can be further fine-tuned to match the initial level of students to avoid frustration for the few for whom activities seemed too difficult. Alternatively, students may be grouped in ways that more advanced students can help students who are still at a beginner's level. With respect to the group dynamics, noting the reported advantages of team

collaboration in gamified experiences (Pazakou et al., 2025), developers of treasure hunts, escape rooms, or other playful activities can take into consideration including more team-building activities to use the opportunity and also increase students' social skills (Georgopoulou, 2024). Future studies may also include (video) observations of the interactions between Game Masters and students to explore the closeness and feedback differences in different games to further explore the possibility that certain activities allow for (slightly) better fulfillment of the student need to feel relatedness and competence.

Conclusions

In conclusion, the study was able to confirm the first research question (RQ1: Do gamified unplugged activities (treasure hunt and escape room) significantly improve college students' computational thinking (CT) knowledge?). There indeed were significant improvements in students' computational thinking knowledge after participating in the summer camp activities. Regarding RQ2 (i.e., How do gender and the type of gamified activity influence computational thinking improvement and students' affective experiences?), specific effects were found for the multiple-choice CT questions. There was a small main effect of gender, with boys generally scoring better than girls. Moreover, while both boys and girls improved in both conditions, the escape room appeared slightly more effective for boys. This led to a significant overall difference in post-test multiple-choice scores, favoring the escape room. However, for the sentence completion and matching tasks, improvement occurred regardless of gender or activity type. With respect to competence beliefs, students in the escape room felt more competent. Further, while boys reported more competence in the escape room, overall, girls from both groups felt more competent compared to boys. With respect to tension, students in the escape room felt less tension. Both boys and girls experienced less tension in the escape room, but boys felt significantly more tense than girls within the escape room condition. Finally, with respect to negative affect, the escape room led to significantly less negative affect. Both boys and girls felt more positively in the escape room, but girls in the escape room felt more positively compared to boys. In the treasure hunt, feelings of boys and girls were comparable.

Finally, to answer the third research question (RQ3: What are students' and game masters' qualitative experiences regarding the computational thinking challenges and the gamification elements of the activities?), the qualitative results were mostly positive. Themes for computational thinking included the activities being engaging and enjoyable, the benefit of breaking down problems into smaller parts, and an expressed interest in further learning. The negative feedback was captured by the themes "losing sight of the bigger picture" (indicating confusion with problem integration) and experiencing "frustration" when tasks were perceived as too difficult. Regarding the experiences with gamification elements, themes included high engagement and motivation, positive experiences with collaborative learning (including team competition), and appreciation for the cognitive challenge presented by the puzzles. Negative feedback was found in "progress-related stress" (especially from competitive elements), challenges with "team dynamics" (unequal contribution, interpersonal friction, skill variability, physical coordination issues), and "unclear learning purpose and outcomes" (students not always understanding the educational objective of the activities).

Notes

1. https://www.twinkl.gr/blog/8-ways-to-take-computing-outdoors, https://www.learning.com/blog/unplugged-programming-activities/.
2. https://www.bebraschallenge.org.

Acknowledgments

The authors wish to express their sincere gratitude to [undisclosed for peer review] for his support and permission to conduct this research during the summer camp program. We extend our heartfelt appreciation to all team leaders who dedicated their time to facilitate the preparation of the activities and the collection of data. Finally, we are especially grateful to all the children who participated in the study and their parents for their trust and support.

Ethical approval

The study was conducted in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments. In accordance with the national research ethics guidelines, the conduct of this study did not require formal ethics committee approval as it consisted only of standard psychometric assessments. Before the commencement of the activities, information about the objectives, the procedures, the data collection methods, and the confidentiality measures was provided to all participants' parents/legal guardians. Written informed consent was obtained from the parents/legal guardians of all participants included in the study.

Disclosure statement

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

Funding

The research received no external funding. All resources and materials were provided through volunteer efforts.

ORCID

Athanasios Christopoulos http://orcid.org/0000-0002-1809-5525
Stylianos Mystakidis http://orcid.org/0000-0002-9162-8340
Chrysostomos Stylios http://orcid.org/0000-0002-2888-6515
Ioannis G. Tsoulos http://orcid.org/0000-0002-2343-2733

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives*. Addison Wesley Longman, Inc.

Belmar, H. (2022). Review on the teaching of programming and computational thinking in the world. *Frontiers in Computer Science*, *4*, 997222. https://doi.org/10.3389/fcomp.2022.997222

Bocconi, S., Chioccariello, 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: State of play and practices from computing education*. European Commission. https://epublications.vu.lt/object/elaba:124209087/124209087.pdf

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, *18*(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 Annual Meeting of the American Educational Research Association* (Vol. 1, pp. 1–25). AERA.

Chen, P., Yang, D., Metwally, A. H. S., Lavonen, J., & Wang, X. (2023). Fostering computational thinking through unplugged activities: A systematic literature review and meta-analysis. *International Journal of STEM Education*, *10*(1), 47. https://doi.org/10.1186/s40594-023-00434-7

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

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Fotaris, P., & Mastoras, T. (2019, October). Escape rooms for learning: A systematic review. In *Proceedings of the European Conference on Games Based Learning* (Vol. 2019, pp. 235–243). https://doi.org/10.34190/GBL.19.179

Georgopoulou, M. S. (2024). Power of synergy: Unlocking the potential of group dynamics through team-building practices in junior high school. *European Journal of Education and Pedagogy*, *5*(2), 12–21. https://doi.org/10.24018/ejedu.2024.5.2.803

Griselda, S. (2024). Gender gap in standardized tests: What are we measuring? *Journal of Economic Behavior & Organization*, *221*, 191–229. https://doi.org/10.1016/j.jebo.2024.03.010

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, *54*, 170–179. https://doi.org/10.1016/j.chb.2015.07.045

Hellín, C. J., Calles-Esteban, F., Valledor, A., Gómez, J., Otón-Tortosa, S., & Tayebi, A. (2023). Enhancing student motivation and engagement through a gamified learning environment. *Sustainability*, *15*(19), 14119. https://doi.org/10.3390/su151914119

Henriksen, D., Henderson, M., Creely, E., Carvalho, A. A., Cernochova, M., Dash, D., Davis, D., & Mishra, P. (2021). Creativity and risk-taking in teaching and learning settings: Insights from six international narratives. *International Journal of Educational Research Open*, *2*, 100024. https://doi.org/10.1016/j.ijedro.2020.100024

Huang, W., & Looi, C. K. (2021). A critical review of literature on "unplugged" pedagogies in K-12 computer science and computational thinking education. *Computer Science Education*, *31*(1), 83–111. https://doi.org/10.1080/08993408.2020.1789411

IJsselsteijn, W. A., de Kort, Y. A. W., & Poels, K. (2013). *The game experience questionnaire*. Technische Universiteit Eindhoven. https://pure.tue.nl/ws/files/21666907/Game_Experience_Questionnaire_English.pdf

Kapur, M. (2008). Productive failure. *Cognition and Instruction*, *26*(3), 379–424. https://doi.org/10.1080/07370000802212669

Kim, H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, *38*(1), 52–54. https://doi.org/10.5395/rde.2013.38.1.52

Krath, J., Schürmann, L., & Von Korflesch, H. F. (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior*, *125*, 106963. https://doi.org/10.1016/j.chb.2021.106963

Kuo, W.-C., & Hsu, T.-C. (2020). Learning computational thinking without a computer: How computational participation happens in a computational thinking board game. *The Asia-Pacific Education Researcher*, *29*(1), 67–83. https://doi.org/10.1007/s40299-019-00479-9

Lampropoulos, G., Keramopoulos, E., Diamantaras, K., & Evangelidis, G. (2023). Integrating augmented reality, gamification, and serious games in computer science education. *Education Sciences*, *13*(6), 618. https://doi.org/10.3390/educsci13060618

Luarn, P., Chen, C. C., & Chiu, Y. P. (2023). Enhancing intrinsic learning motivation through gamification: A self-determination theory perspective. *The International Journal of Information and Learning Technology*, *40*(5), 413–424. https://doi.org/10.1108/IJILT-07-2022-0145

Martia, L. A., & Nurafni, N. (2023). The impact of treasure hunt games on learning outcomes in time measurement. *Al-Jabar: Jurnal Pendidikan Matematika*, *14*(2), 515–525. https://doi.org/10.24042/ajpm.v14i2.18237

Mystakidis, S., Berki, E., & Valtanen, J.-P. (2021). Deep and Meaningful E-Learning with Social Virtual Reality Environments in Higher Education: A Systematic Literature Review. *Applied Sciences*, *11*(5), 2412 https://doi.org/10.3390/app11052412

Pardhan, S., Parkin, J., Trott, M., & Driscoll, R. (2022). Risks of digital screen time and recommendations for mitigating adverse outcomes in children and adolescents. *The Journal of School Health*, *92*(8), 765–773. https://doi.org/10.1111/josh.13170

Pazakou, A.-M., Mystakidis, S., & Kazanidis, I. (2025). Collaboration in a Virtual Reality Serious Escape Room in the Metaverse Improves Academic Performance and Learners' Experiences. *Future Internet*, *17*(1), 21 https://doi.org/10.3390/fi17010021

Rocka, A., Jasielska, F., Madras, D., Krawiec, P., & Pac-Kożuchowska, E. (2022). The impact of digital screen time on dietary habits and physical activity in children and adolescents. *Nutrients*, *14*(14), 2985. https://doi.org/10.3390/nu14142985

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Santos, R. M. S., Mendes, C. G., Sen Bressani, G. Y., de Alcantara Ventura, S., de Almeida Nogueira, Y. J., de Miranda, D. M., & Romano-Silva, M. A. (2023). The associations between screen time and mental health in adolescents: A systematic review. *BMC Psychology*, *11*(1), 127. https://doi.org/10.1186/s40359-023-01166-7

Sharpe, D., & Cribbie, R. A. (2023). Analysis of treatment-control pre-post-follow-up design data. *The Quantitative Methods for Psychology*, *19*(1), 25–46. https://doi.org/10.20982/tqmp.19.1.p025

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

Snelson, C. (2022). Quest-based learning: A scoping review of the research literature. *TechTrends*, *66*(2), 287–297. https://doi.org/10.1007/s11528-021-00674-w

Sun, D., Zhu, C., Xu, F., Li, Y., Ouyang, F., & Cheng, M. (2024). Transitioning from introductory to professional programming in secondary education: Comparing learners' computational thinking skills, behaviors, and attitudes. *Journal of Educational Computing Research*, *62*(3), 427–454. https://doi.org/10.1177/07356331231204653

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

Torres-Torres, Y. D., Román-González, M., & Perez-Gonzalez, J. C. (2024). Didactic strategies for the education of computational thinking from a gender perspective: A systematic review. *European Journal of Education*, *59*(2), e12640. https://doi.org/10.1111/ejed.12640

Triantafyllou, S. A., Sapounidis, T., & Farhaoui, Y. (2024a). Gamification and computational thinking in education: A systematic literature

review. *Salud, Ciencia y Tecnologia-Serie*, *3*, 659. https://doi.org/10.56294/sctconf2024659

Triantafyllou, S. A., Sapounidis, T., & Farhaoui, Y. (2024b). Students' development of computational thinking and teachers' professional development via bebras through gamification: A systematic literature review. https://doi.org/10.56294/dm2024.582

Veldkamp, A., Knippels, M. C. P., & van Joolingen, W. R. (2021). Beyond the early adopters: Escape rooms in science education. *Frontiers in Education*, *6*, 622860. https://doi.org/10.3389/feduc.2021.622860

Vonitsanos, G., Moustaka, I., Doukakis, S., & Mylonas, P. (2024, May). *Transforming education in the digital age: Exploring the dimensions of Education 4.0* [Paper presentation]. 2024 IEEE Global Engineering Education Conference (EDUCON) (pp. 1–10). IEEE. https://doi.org/10.1109/EDUCON60312.2024.10578747

Zahedi, L., Batten, J., Ross, M., Potvin, G., Damas, S., Clarke, P., & Davis, D. (2021). Gamification in education: A mixed-methods study of gender on computer science students' academic performance and identity development. *Journal of Computing in Higher Education*, *33*(2), 441–474. https://doi.org/10.1007/s12528-021-09271-5

Appendix —Instruments

All the instruments were provided in participants' native language.

Demographics

Cabin Nr.://Gender: Boy, Girl, Prefer not to answer//**Age**: 13, 14, 15

Knowledge assessment

Part A—Multiple choice questions

1. **What is Computational Thinking?**

   a. It is the ability to translate complex problems into a language that computers can understand.
   b. **It is the process of thinking about problems and solutions in a way that a computer can help to solve them.**
   c. It is the method to make computers think like humans.
   d. It is the concept of programming computers to solve problems.

2. What does the term "algorithm" mean in Computer Science?

   a. It is a set of instructions designed to perform a specific task.
   b. It is a type of data structure that stores information.
   c. It is a programming language.
   d. It is a term used for the process of debugging code.

3. **What does "Algorithmic Thinking" involve in Computational Thinking?**

   a. Using algorithms to make a computer think like a human.
   b. **Creating a set of step-by-step instructions to solve a problem.**
   c. Breaking down a problem into smaller components.
   d. Eliminating unnecessary details from a problem.

4. **What is the purpose of Decomposition in Computational Thinking?**

   a. To write a computer program.
   b. To compile and run a computer program.
   c. **To break down a complex problem into smaller, manageable parts.**
   d. To identify patterns and trends in a set of data.

5. **What do we mean by "Pattern Recognition" in Computational Thinking?**

   a. Identifying the frequency of a certain pattern in a computer program.
   b. **Identifying similarities or patterns among small, decomposed problems to aid in solving larger ones.**
   c. Using a particular pattern to design a computer system.
   d. Recognizing the pattern in which a computer program executes.

6. **What does "Abstraction" represent in Computational Thinking?**

   a. It is a programming language.
   b. **It is a way to simplify complex systems by ignoring or hiding some details while focusing on the important ones.**
   c. It is a type of data structure.
   d. It is a process to debug code.

7. **What is a "binary system"?**

   a. A system of numbers using only the digits 1 and 0.
   b. A system of two computers connected over a network.
   c. A system that can only perform two operations.
   d. A system that can only handle two types of data.

8. **What is "debugging" in the context of Computer Science?**

   a. It is the process of identifying and removing errors in code.
   b. It is the process of enhancing the efficiency of an algorithm.
   c. It is a programming language.
   d. It is a method of writing a computer program.

9. **What is a "pseudo-code"?**

   a. It is a type of programming language.
   b. **It is a simple way of expressing a program or algorithm that uses a mixture of English phrases and syntax resembling programming languages.**
   c. It is a form of shorthand notation for representing binary data.
   d. It is a type of code used for debugging purposes.

10. **What is "Parallel Processing" in Computational Thinking?**

a. It is a process of breaking down a large problem into smaller parts.
b. **It is a method of executing multiple tasks simultaneously.**
c. It is a way to debug multiple codes at the same time.
d. It is a type of data structure used for storing large amounts of data.

Part B—Fill in the blanks

[abstraction, array, Boolean, database, debugging, exception handling, hashing, If, linked list, list, polymorphism, recursion, sort(), variables, while]

1. A _____ loop will continue to execute until the condition is false. [while]
2. The _____ function is used to sort elements in a list. [sort ()]
3. The purpose of _____ in a program is to store data that can be used by the program. [variables]
4. A _____ is a type of data that can only be True or False. [Boolean]
5. A _____ statement is used to make a decision based on a condition. [if]
6. A _____ is a set of related values that can be numeric or string, among others. [array]
7. A _____ is a data type that can store a sequence of elements. [list]
8. The process of checking for errors in a program is known as _____. [debugging]
9. A _____ is a model that allows data to be stored, manipulated, and retrieved. [database]
10. In computational thinking, _____ is a way to simplify complex systems by ignoring or hiding some details. [abstraction]

Part C—Matching

Match the following computational thinking concepts with their definitions:

1. Decomposition	A. Breaking down a complex problem into manageable parts.
2. Pattern recognition	B. Identifying similarities or shared characteristics among data, information, or scenarios.
3. Abstraction	C. Simplifying complex systems by hiding some details.
4. Algorithm design	D. Developing a step-by-step solution for a problem.
5. Binary system	E. A system of numbers using only the digits 1 and 0.
6. Debugging	F. The process of identifying and removing errors in a program.
7. Pseudo-code	G. A simple way of expressing a program or algorithm that uses a mixture of English phrases and syntax resembling programming languages.
8. Heuristics	H. An approach to problem solving which provides a good enough solution under given circumstances.
9. Parallel processing	I. The simultaneous execution of the same task split into independent parts on multiple processors in a computer.
10. Encapsulation	j. The bundling of data and methods that work on that data within one unit, often used in object-oriented programming.

Survey

Please indicate how you felt while playing the game for each of the items, on the following scale: 0 = not at all, 1 = slightly, 2 = moderately, 3 = fairly, 4 = extremely

C1. Competence beliefs

1. I felt skillful.
2. I felt competent.
3. I was good at it.
4. I felt successful.
5. I was fast at reaching the game's targets.

C2. Tension/annoyance

1. I felt annoyed.
2. I felt irritable.
3. I felt frustrated.

C3. Negative affect

1. It gave me a bad mood.
2. I thought about other things.
3. I found it tiresome.
4. I felt bored.

Discussion and reflection

Topic: Computational thinking

Think back to the escape room/treasure hunt activity. Describe a time when you had to break down a complex puzzle or challenge into smaller, easier-to-solve steps. How did this approach help you solve the puzzle? Did you use any strategies that you think a computer programmer might also use?

Topic: Games

What aspects of the escape room/treasure hunt experience did you find the most fun or engaging? Did the competitive elements, puzzles, teamwork, or the story of the game make it more enjoyable? If you could design your own escape room or game, what features would you include to make it even more fun and interesting?