Computational Thinking as a Cognitive Construct

Cognitive Correlates, Assessment & Curriculum Design

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Katerina (Aikaterini) Tsarava

aus Thessaloniki, Griechenland

Tübingen

2020

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	22.02.2021
Stellvertreter Dekan:	Prof. Dr. József Fortágh
1. Berichterstatter:	Prof. Dr. Martin Butz
2. Berichterstatter:	Prof. Dr. Hartmut Leuthold

“Computational Thinking is a universal metaphor […]

of reasoning used by both mankind and machines.”

-Jeanette Wing

Abstract

Computational thinking (CT) has been coined a fundamental 21st skill comparable to literacy and numeracy. As a cognitive ability underlying programming and coding skills, CT has been suggested to be fostered early on in education. Accordingly, the last decade significant research effort has been devoted to developing educational activities for CT, integrated into formal and informal educational settings. However, despite the various research on CT, its definition and the respective assessment approaches are still in their infancy. Consequently, the lack of a consensus definition of CT and the limited validated assessment tools for measuring CT restrain the empirical evaluation of the proposed educational materials.

This thesis aims to investigate CT as a cognitive construct and provide an evidence-based definition for it, focusing on the underinvestigated population of elementary school children. To achieve this, first, a review of the literature was conducted to identify concrete CT concepts. Accordingly, a CT curriculum for 3rd and 4th graders was developed, taking into consideration the complexity of the CT concepts and therefore integrating game-based and embodied activities to provide a low-threshold introduction to CT. The novel parts of the curriculum, which are a series of life-size board games, were iteratively evaluated for their usability with adults and children before being integrated into it. After improvements to the games, the curriculum underwent a pilot and an effectiveness evaluation, the latter one designed as a randomized control field trial. In order to measure the effectiveness of the curriculum, a CT assessment tool was developed. Additionally, a correlational analysis was performed in both evaluation phases in order to investigate associations of CT with other cognitive abilities, and therefore complement the nomological network of CT and define the construct.

Results on the effectiveness of the curriculum showed positive effects on students' CT abilities and therefore appropriateness of the curriculum design, the development and evaluation procedures followed. Moreover, the proposed CT assessment seems reliable for measuring CT at elementary school students and can be used in future studies. Results on the cognitive correlates of CT revealed positive associations of CT with verbal reasoning-, non-verbal visuospatial-, and complex numerical abilities. These results, compared to similar research in other age groups, show similarities but also differences. This implies that CT development is supported by different cognitive abilities across age groups. Additionally, the cognitive

abilities investigated in this research could only partially explain CT. This further supports the argument that CT is a specific cognitive ability that builds on and recruits a convolute of several other cognitive abilities, which are not yet extensively investigated in relation to CT.

Zusammenfassung

Computational Thinking (CT) Computational Thinking (CT) [Informatisches Denken] ist eine essentielle Fähigkeit des 21. Jahrhunderts, die von Ihrer gesellschaftlichen Bedeutung mit Lesen, Schreiben und Rechnen vergleichbar ist. Diese kognitive Fähigkeit bildet eine Grundlage für Programmier- und Kodierfähigkeiten. Damit ist sie heutzutage zentral für die Ausübung vieler Berufe. Aufgrund dessen gibt es Bestrebungen, CT frühzeitig während der Schulbildung zu fördern. Dementsprechend wurden in den letzten zehn Jahren erhebliche Anstrengungen unternommen, um Bildungsaktivitäten für CT zu entwickeln, die in formelle und informelle Bildungssituationen integriert werden können. Trotz der verschiedenen Forschungen zu CT stecken seine Definition und die entsprechenden Bewertungsansätze noch in den Kinderschuhen. Das Fehlen einer einheitlichen Definition sowie die nur begrenzt validierten Bewertungsinstrumente zur Messung von CT schränken die empirische Evaluation der vorgeschlagenen Lehrmaterialien folglich ein.

Das Ziel dieser Arbeit ist es, CT als kognitives Konstrukt zu untersuchen und eine evidenzbasierte Definition dafür zu liefern. Hierbei liegt der Schwerpunkt auf der noch wenig untersuchten Population von Grundschulkindern. In einem ersten Schritt wurde eine Literaturanalyse durchgeführt, um konkrete CT-Konzepte zu identifizieren. Darauf aufbauend wurde ein CT-Lehrplan für Schüler*nnen der 3. und 4. Klasse entwickelt, der die Komplexität der CT-Konzepte berücksichtigt und spielbasierte und verkörperte („embodied“) Aktivitäten integriert, um eine einfache Einführung in CT zu ermöglichen. Hierbei handelt es sich um eine Reihe lebensgroßer Brettspiele. Diese wurden vor ihrer Integration iterativ auf ihre Verwendbarkeit mit Erwachsenen und Kindern bewertet. Nach Verbesserungen der Spiele wurde der Lehrplan einer Pilot- und Effektivitätsbewertung unterzogen, wobei letztere als randomisierte Kontrollfeldstudie konzipiert wurde. Um die Wirksamkeit des Lehrplans zu testen, wurde ein CT-Bewertungsinstrument entwickelt. Zusätzlich wurde in beiden Bewertungsphasen eine Korrelationsanalyse durchgeführt, um Assoziationen von CT mit anderen kognitiven Fähigkeiten zu untersuchen und damit das nomologische Netzwerk von CT zu ergänzen und das Konstrukt zu definieren.

Die Ergebnisse zur Wirksamkeit des Lehrplans zeigten positive Auswirkungen auf die CT-Fähigkeiten der Schüler*nnen und damit auf die Angemessenheit des Lehrplandesigns. Darüber hinaus scheint die vorgeschlagene CT-Bewertung für die Messung von CT bei

Grundschüler*innen zuverlässig zu sein und kann in zukünftigen Studien verwendet werden. Die Ergebnisse zu den kognitiven Korrelaten von CT zeigten positive Assoziationen von CT mit verbalen Argumentations-, nonverbalen visuellen und komplexen numerischen Fähigkeiten. Der Vergleich dieser Ergebnisse mit bereits existierenden Studien in anderen Altersgruppen zeigt viele Ähnlichkeiten, aber auch Unterschiede. Dies impliziert, dass die CT-Entwicklung durch unterschiedliche kognitive Fähigkeiten in verschiedenen Altersgruppen unterstützt wird. Darüber hinaus konnten die in dieser Studie untersuchten kognitiven Fähigkeiten CT nur teilweise erklären. Dies stützt ferner das Argument, dass CT eine spezifische kognitive Fähigkeit ist, die auf mehreren anderen kognitiven Fähigkeiten aufbaut und diese rekrutiert, was im Bezug auf CT noch nicht umfassend untersucht wurde.

Table of Contents

List of Abbreviations

CS Computer Science

CT Computational Thinking

CTt Computational Thinking test

ICT Information and Communication Technologies

OER Open Educational Resources

OSF Open Science Framework

STEM Science Technology Engineering Mathematics

List of Figures

List of Tables

List of Publications

In fulfilment of the requirements of this cumulative dissertation, six articles and manuscripts (under review or in preparation) are submitted. The contribution of all authors is given according to CRediT[1]. As part of my PhD research, I also co-authored a detailed course manual[2], described in Studies 4 and 6 of this dissertation, which is not included in this dissertation for space convenience.

Section 6: Computational Thinking: Curriculum Design & Evaluation

Study 1: **Tsarava**, **K**., Moeller, K., Pinkwart, N., Butz, M., Trautwein, U., & Ninaus, M. (2017). **Training computational thinking: Game-based unplugged and plugged-in activities in primary school.** Proceedings of the *11th European Conference on Game-Based Learning,* 687-695. Reading, UK: Academic Conferences and Publishing International Limited.[3]

CRediT author statement

Katerina Tsarava: Conceptualization, Investigation, Resources, Writing - Original Draft, Writing - Review & Editing, Visualization; Korbinian Moeller: Conceptualization, Writing - Review & Editing, Supervision, Funding Acquisition; Niels Pinkwart: Funding Acquisition, Writing - Review & Editing; Martin V. Butz: Funding Acquisition, Writing - Review & Editing; Ulrich Trautwein: Writing - Review & Editing, Supervision, Funding Acquisition; Manuel Ninaus: Conceptualization, Writing - Review & Editing, Supervision, Funding Acquisition.

Study 2: **Tsarava**, **K**., Moeller, K., & Ninaus, M. (2018). **Training computational thinking through board games: The case of Crabs & Turtles**. *International Journal of Serious Games, 5(2)*, 25-44.[4]

CRediT author statement

---

[1] CRediT (Contributor Roles Taxonomy) author statement. https://casrai.org/credit/

[2] Leifheit, L., **Tsarava, K.**, Ninaus, M., & Moeller, K. (2018). *"Verstehen wie Computer denken" - Ein Trainingsprogramm zur Förderung von informatischem Denken und systematischen Problemlösefähigkeiten besonders begabter Kinder im Grundschulalter.* Reihe Hector Core Courses.

[3] Reprinted by permission of the *Academic Conferences and Publishing International Limited*. Original reference: Tsarava, K., Moeller, K., Pinkwart, N., Butz, M., Trautwein, U., & Ninaus, M. (2017). Training computational thinking: Game-based unplugged and plugged-in activities in primary school. *Proceedings of the 11th European Conference on Game Based Learning (pp. 687-695).* Reading, UK: Academic Conferences and Publishing International Limited.

[4] Reprinted by permission of the *International Journal of Serious Games*. Original reference: Tsarava, K., Moeller, K., & Ninaus, M. (2018). Training Computational Thinking through board games: The case of Crabs & Turtles. *International Journal of Serious Games, 5(2)*, 25-44. https://dx.doi.org/10.17083/ijsg.v5i2.248

Katerina Tsarava: Conceptualization, Data curation, Investigation, Resources, Writing - Original Draft, Writing - Review & Editing, Visualization; Korbinian Moeller: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding Acquisition; Manuel Ninaus: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding Acquisition.

Study 3: **Tsarava**, **K**., Moeller, K., & Ninaus, M. (2019). **Board Games for Training Computational Thinking.** Proceedings of *Games and Learning Alliance conference (GALA 2018) - Lecture Notes in Computer Science*. Springer.[5]

CRediT author statement

Katerina Tsarava: Conceptualization, Data curation, Investigation, Resources, Writing - Original Draft, Writing - Review & Editing, Visualization; Korbinian Moeller: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding Acquisition; Manuel Ninaus: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Funding Acquisition.

Section 7: Computational Thinking: Cognitive Definition & Assessment

Study 4: **Tsarava**, **K**., Leifheit, L., Ninaus, M., Román-González, M., Butz, M. V., Golle, J., Trautwein, U., & Moeller, K. (2019). **Cognitive Correlates of Computational Thinking: Evaluation of a Blended Unplugged/Plugged-In Course.** *Proceedings of WiPSCE '19: Workshop in Primary and Secondary Computing Education (WiPSCE '19).* New York, NY, USA: ACM. [6]

CRediT author statement

Katerina Tsarava: Conceptualization, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project Administration; Luzia Leifheit: Investigation, Writing - Review & Editing, Project Administration; Manuel Ninaus: Conceptualization, Writing - Review & Editing, Supervision, Funding Acquisition; Marcos Roman-Gonzalez: Methodology, Writing - Review & Editing; Martin V. Butz: Writing - Review & Editing, Funding Acquisition, Supervision; Jessika Golle: Writing - Review & Editing, Funding Acquisition; Ulrich Trautwein: Supervision, Funding Acquisition; Korbinian Moeller: Conceptualization, Writing - Review & Editing, Supervision, Funding Acquisition.

[5] Reprinted by permission of the *Springer Nature*. Original reference: Tsarava K., Moeller K., Ninaus M. (2019). Board Games for Training Computational Thinking. In: Gentile M., Allegra M., Söbke H. (eds) Games and Learning Alliance. GALA 2018. Lecture Notes in Computer Science, vol 11385. Springer, Cham. https://doi.org/10.1007/978-3-030-11548-7_9

[6] Reprinted by permission of the *Association for Computer Machinery (ACM)*. Original reference: Tsarava, K., Leifheit, L., Ninaus, M., Román-González, M., Butz, M. V., Golle, J., Trautwein, U., & Moeller, K. (2019). Cognitive Correlates of Computational Thinking: Evaluation of a Blended Unplugged/Plugged-In Course. *Proceedings of WiPSCE '19: Workshop in primary and Secondary Computing Education (WiPSCE '19)* (Article No. 24). New York, NY, USA: ACM. https://dx.doi.org/10.1145/3361721.3361729

Study 5: **Tsarava**, **K**., Moeller, K., Román-González, M., Golle, J., Leifheit, L., Butz, M. V., & Ninaus, M. (under review). **A Cognitive Approach to Defining and Assessing Computational Thinking: An Empirical Study in Primary School**.[7]

CRediT author statement

Katerina Tsarava: Conceptualization, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project Administration; Korbinian Moeller: Conceptualization, Writing - Review & Editing, Supervision, Funding Acquisition; Marcos Roman-Gonzalez: Writing - Review & Editing; Jessika Golle: Writing - Review & Editing, Funding Acquisition; Luzia Leifheit: Investigation, Writing - Review & Editing, Project Administration; Martin V. Butz: Writing - Review & Editing, Funding Acquisition, Supervision; Manuel Ninaus: Conceptualization, Writing - Review & Editing, Supervision, Funding Acquisition.

Study 6: **Tsarava**, **K**., Leifheit, L., Ninaus, M., Golle, J., Trautwein, U., & Moeller, K. (in preparation). **Evaluation of a Computational Thinking Intervention for Elementary School Children: A Randomized Controlled Field Trial.**

CRediT author statement

Katerina Tsarava: Conceptualization, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project Administration; Luzia Leifheit: Investigation, Writing - Review & Editing, Project Administration; Manuel Ninaus: Conceptualization, Writing - Review & Editing, Supervision, Funding Acquisition; Jessika Golle: Writing - Review & Editing, Funding Acquisition; Ulrich Trautwein: Supervision, Funding Acquisition; Korbinian Moeller: Conceptualization, Writing - Review & Editing, Supervision, Funding Acquisition.

[7] This manuscript is submitted at the journal *Computers & Education* and is in the review process (by December 2020).

PART I: GENERAL INTRODUCTION

1 Computational Thinking: Definition

Computational thinking (CT) is a term attracting considerable and increasing educational and research interest over the past few decades. The initial articulation of the term is credited to Seymour Papert (for a historical overview on the term conceptualization, see Kong & Abelson, 2019) and the foundations of constructionism (Papert & Harel, 1991). In 2006 though, Janette Wing launched the starting signal for the development of a whole new field of research in the interdisciplinary field bounded mainly within the domains of educational sciences, computer science (CS), psychology, and other complementary to these science domains. This first conceptualization of the term described CT as a way of *"solving problems, designing systems, and understanding human behavior, by drawing on the concepts fundamental to computer science (...) Computational thinking is reformulating a seemingly difficult problem into one we know how to solve, perhaps by reduction, embedding, transformation, or simulatio*n." (Wing, 2006, p. 33). CT already in this initial attempt of definition was introduced as a fundamental skill distinctly different from programming, that requires thinking in higher levels of abstraction, draws upon mathematical and engineering thinking and is essential for everyone and not just computer scientists (Armoni, 2016; Settle et al., 2013).

As already indicated in the first conceptualization of the term, CT is "a fundamental, not rote skill" (Wing, 2006, p. 35). This clear distinction emphasizes the difference of CT from computer programming, often referred also as coding. In CS, programming and coding refer to more practical and to a lesser extend cognitive skills, describing the acts of developing computer programs and writing computer code respectively. However, these two terms are very often referenced as related or complementary to CT (Armoni, 2016). Central concepts in coding and computer programming are the notions of sequencing, loops, parallelism, events, conditionals, operators and data (variables and constants) (Brennan & Resnick, 2012a). CT draws on cognitive processes such as algorithmic thinking, conditional logic, decomposition, abstraction, pattern matching, parallelization, and evaluation (e.g., Astrachan & Briggs, 2012; Wing, 2010). These cognitive processes can reflect concepts fundamental to coding and computer programming. However, though driven from CS in general and programming or coding in particular, they can be applied to various other domains, and in real-life problems or activities as well (P. S. Wang, 2015). Therefore, though CT seems a prerequisite for developing coding and programming skills, CT is considered a broader cognitive skill. As such, and also as a universal problem-solving skill CT is related to the development of cognitive skills in various contexts, beyond CS (Armoni, 2016; Moreno-Leon et al., 2018; Yaşar, 2018a).

As a cognitive ability underpinning programming and coding skills, CT has been suggested a competence comparable to literacy and numeracy, to be acquired early on in education (e.g., Wing, 2006a; Yadav et al., 2014). Consequently, over the last decade CT has been attracting increasing research interest (K. Tang, 2019), being considered a *21st-century* and a *universal skill* (e.g., D. Barr et al., 2011; Settle et al., 2013; J. Voogt et al., 2013; Yadav et al., 2011). This interest has inspired multiple efforts for integrating CT in school curricula either as a standalone topic or interdisciplinarily in the curricula of other STEM domains such as CS, but also beyond that (Chiprianov & Gallon, 2016; Lockwood & Mooney, 2017; Moreno-Leon et al., 2018; Settle et al., 2013; Weintrop et al., 2016). This integration has been observed at all levels of education, from pre-school to university level, both in formal and informal educational settings (K. Tang, 2019).

The increasing interest of integrating CT in school curricula and the consequent development of numerous educational materials for fostering CT demands appropriate assessment tools for evaluating their effectiveness. The design and validation of reliable CT assessments, however, demands a well-defined construct to be the targeted construct of assessment. There have been different efforts to elaborate on a purposeful definition of CT over the last years (Garcia-Peñalvo, 2016; Grover & Pea, 2013; Lockwood & Mooney, 2017; Yaşar, 2018b), but no consensus definition has been formulated yet in an evidence-based manner and therefore widely popularized. As a working definition and conceptualization of CT for this work, I used a definition that resulted from a review of the literature by Shute et al. (2017). This review describes CT as *"the conceptual foundation required to solve problems effectively and efficiently (i.e., algorithmically, with or without the assistance of computers) with solutions that are reusable in different contexts"* (Shute et al., 2017, p.142). This interpretation is in line with my approach of defining CT as the cognitive underpinning of computer programming, that goes beyond CS and can find application as a cognitive strategy in various other aspects of life.

In the following four sections, first, a cognitive interpretation of CT is attempted (section 2), then, a description of integrating CT into a curriculum is presented (section 3), the CT assessment is discussed (section 4), and last, the objectives of this dissertation are introduced (section 5).

2 Computational Thinking: Relevance to Cognition

A recent summary of results from various empirical studies has shown multiple transfer effects of programming - as an integrative approach of teaching, learning, and assessing CT- on cognition (Scherer et al., 2019). This meta-analysis demonstrated positive near and far transfer effects on situations requiring creative thinking, mathematical skills, metacognition, spatial skills, and reasoning. Even though the most considerable part of research on potential transfer effects of programming is focused on the elementary school level, studies investigating the cognitive underpinnings of CT consider students at university level and, to a smaller extent, high school students (Román-González et al., 2018). Apart from the lack of studies on the cognitive correlates of CT in different age groups, the lack of a precise positioning of CT within the nomological network of cognitive abilities was also supported by a recent review on the empirical assessment of CT (X. Tang et al., 2020). Although 80 per cent of the reviewed studies on CT assessment measured cognitive constructs, the cognitive definition of CT is not yet well-grounded independently of the domain of programming skills.

In the following sections, the available empirical evidence on associations of CT with other cognitive abilities at pre-school, elementary, middle, high-school, and university level is presented. These associations, in many cases, are investigated within the context of programming or other STEM activities, a research limitation already been pointed out in literature (e.g., Scherer et al., 2019; Tang et al., 2020). The conclusions drawn by examination of the current state of the literature posit that there are specific associations of CT (and, therefore, programming skills) with other cognitive abilities. Nevertheless, these associations should be further investigated across age groups and with larger sample sizes in order to gain a better understanding of how this construct develops.

2.1 CT and numerical/mathematical cognition

The association between programming and numerical/mathematical abilities was already reported almost thirty years ago (e.g., Pea & Kurland, 1984). A significant number of subsequent empirical studies substantiated these results (e.g., Bergin & Reilly, 2006; Byrne & Lyons, 2001; McCoy & Burton, 1988; Nowaczyk, 1983). CT literature so far points towards programming as the most common way of fostering CT in the early years of school (Moreno-Leon et al., 2018). Consequently, investigating associations of CT with numerical

/mathematical abilities seems relevant for an initial understanding of CT as a unique cognitive ability, which could potentially find application beyond CS (Settle et al., 2013).

From a theoretical perspective, the associations between CT and numerical/mathematical abilities have been visually portrayed by Sneider et al. (2014) with a Venn diagram that indicated besides unique capabilities of mathematical and computational thinking; capabilities considered common for both (see Figure 1). The similarities between the two support the notion that CT could be used as a teaching approach infused in different STEM domain classes (e.g., Perkovic et al., 2010; Settle et al., 2012, 2013; Weintrop et al., 2016).

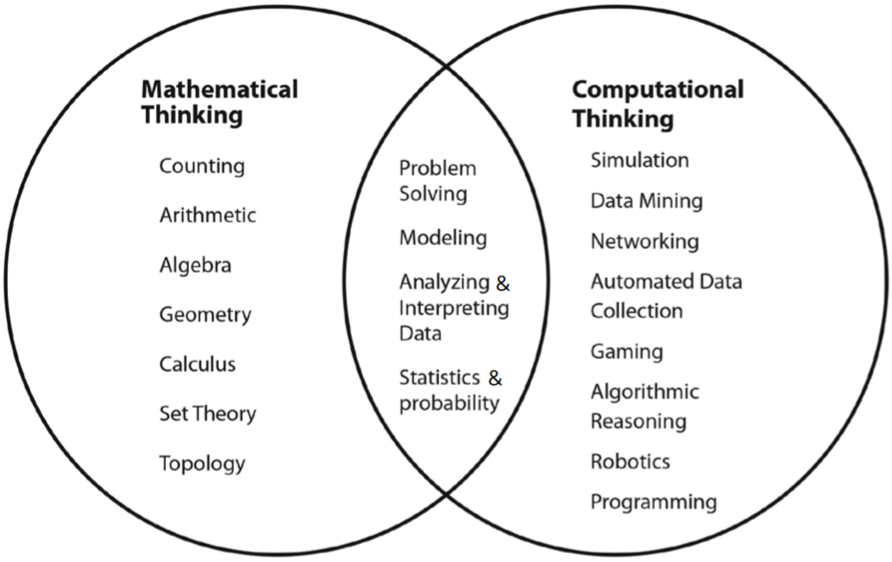

Figure 1. Venn diagram of mathematical and computational thinking (Sneider et al., 2014).

From an empirical perspective, associations between CT and numerical/mathematical cognition have been inconsistent among studies and across age groups (for a brief review, see *Study 5: A cognitive approach to defining and assessing computational thinking: An empirical study in primary school*). There have been studies showing no significant association between CT (or programming) and numerical/mathematical abilities (Ambrosio et al., 2014; Román-González et al., 2017) and others that indicated positive weak to high associations (Prat et al., 2020; Román-González et al., 2018; Werner, 2020).

The inconsistent results on the association of CT and programming skills with numerical/mathematical cognition follow a pattern that indicates that the association seems more evident in younger than in older populations. This pattern may be explained by the fact that numerical/mathematical abilities may be a prerequisite for thinking computationally at an early stage of cognitive development rather than that they are later on when a specific threshold of numerical/mathematical abilities has been reached or exceeded (Tsarava et al., 2019).

2.2 CT and language ability

The association between CT and language abilities has been investigated in quite a few studies so far (for a brief review, see *Study 5: A cognitive approach to defining and assessing computational thinking: An empirical study in primary school*); however, these studies can be found less frequently than similar ones investigating the association of CT with other cognitive abilities. Significant positive associations have been reported in populations of pre-schoolers (Marinus et al., 2018), middle school children (Howland & Good, 2015), high-school students (Román-González et al., 2017), and adults (Prat et al., 2020). Interestingly, one of these studies found a positive correlation between CT and verbal ability, while did not find a significant one between CT and numerical/mathematical abilities (Román-González et al., 2017).

These results seem to indicate that language ability is relevant for CT and programming skills. In some cases, this relevance was even more pronounced than the one of numerical/mathematical abilities.

2.3 CT and visuospatial abilities

Several empirical studies have supported the association between CT (and programming skills) and visuospatial abilities over the last decade (for a brief review, see *Study 5: A cognitive approach to defining and assessing computational thinking: An empirical study in primary school*). These studies provide evidence on the association across different age groups, like elementary school children (e.g., Città et al., 2019), middle school children (e.g., Román-González et al., 2017), university students (e.g., Jones & Burnett, 2008), and adults (e.g., Parkinson & Cutts, 2019).

In brief, these studies suggest that visuospatial abilities are significantly associated with CT, and the results have been consistent among age groups, from early elementary school students up to adults.

2.4 CT and general cognitive ability

CT is often described in the literature as a problem-solving process (for a review, see Kalelioğlu et al., 2016). Problem-solving skills are closely related to specific aspects of fluid intelligence

(Carroll, 1993). According to various empirical studies, the ability to program is also associated with non-verbal and general intelligence. In comparison, the empirical evidence on the association of CT with problem-solving ability has been evaluated explicitly in very few studies (Román-González et al., 2017).

Having already described CT as a cognitive ability that is closely related to programming skills and thus, having strong indications of the association between both CT and programming ability with problem-solving skills, it seems we can presume and expect an association between CT and cognitive abilities in general.

3 Computational Thinking: Curricula

During the past few years, CT has been systematically integrated into the official curricula of many countries (e.g., for the UK, see Brown et al., 2014; for France, see Chiprianov & Gallon, 2016; for Australia, see Falkner et al., 2014; for North Macedonia, see Jovanov et al., 2016; for Taiwan, see S. Kong & Abelson, 2019). The integration has occurred with CT fostered either as a standalone teaching subject within CS curricula or as an interdisciplinary teaching approach of other STEM-related teaching subjects (e.g., V. Barr & Stephenson, 2011; Dierbach et al., 2011; Weintrop et al., 2016).

Apart from the consequent integration of CT in university-level education (e.g., Dierbach et al., 2011) where the transfer of knowledgable workers to the technology-oriented job market is immediate, the efforts of promoting CT are also focusing on the secondary educational level (e.g., Settle et al., 2012). Lately, such efforts have been embraced in elementary school education (e.g., Brackmann et al., 2017) as well and gradually, are also being experimentally introduced in pre-school settings (e.g., Sullivan et al., 2013). These efforts are observed both in formal and informal educational settings worldwide.

Though CT still lacks an elaborated definition, its educational value is broadly supported by the research community, educational policymakers, professional associations and non-governmental initiatives worldwide (e.g., Qualls & Sherrell, 2010; for a short review, see also *Study 6: Evaluation of a Computational Thinking Intervention for Elementary School Children: A Randomized Controlled Field Trial).* CT as an integral part of ICT literacy has caused a transparent shift in education by switching the main focus from teaching particular technological tools to inspiring students' understanding of how technology works and trigger students' technological creations (Curzon et al., 2014). In other words, CT provides the pathway for students to not just behave as consumers of technology but be also the potential creators of technology (i.e., prosumers).

The continually increasing interest in fostering CT has resulted in plentiful educational CT materials. Accordingly, there have been various frameworks proposed for designing CT curricula to foster CT as a broader cognitive ability applied to different courses and in different contexts (e.g., Curzon et al., 2014; Perković et al., 2010). The effectiveness of these materials (either as complete curricula or independent activities) has been explored in several studies (e.g., Aggarwal et al., 2017; Tran, 2019; Van Dyne & Braun, 2014; for a short review, see also *Study 6: Evaluation of a Computational Thinking Intervention for Elementary School Children: A Randomized Controlled Field Trial*). Regardless of the various studies on the effectiveness of CT curricula, very few of them manage to report relevant statistical information (i.e., effect sizes, confidence intervals, and levels) about the learning effects of CT interventions (McGill & Decker, 2020). This gap in research does not allow for generalization about the effectiveness of the teaching approaches for CT followed so far. Consequently, the lack of adequate statistical results of empirical research on the effectiveness of CT curricula delays the advancement of the research field.

To conclude, at the same time, that plenty of CT curricula and other educational materials have been designed and implemented worldwide, their effectiveness has not yet been thoroughly investigated. This can be partly explained by the lack of reliable CT assessment tools. The assessment instruments for measuring CT and therefore, the learning outcomes of these educational materials are still in an early stage of development and validation (Román-González et al., 2019; X. Tang et al., 2020). In the following section, an overview of the existing research effort for the development and validation of such tools is presented.

4 Computational Thinking: Assessment

The evaluation of the numerous curricula and educational materials designed to foster CT is highly dependent on appropriate CT assessment tools. These measurement instruments are required in order to assess prior CT abilities effectively, monitor the learning development of CT, and measure learning outcomes. The research on this field has progressed a lot in the last few years; however, there is still a lot to be done for the creation and validation of assessment tools that allow for reliably assessing CT development across age groups (for a brief review, see *Study 4: Cognitive Correlates of Computational Thinking: Evaluation of a Blended Unplugged/Plugged-In Course*).

The most recent review of studies on the assessment of CT has revealed multiple weaknesses in the empirical research conducted so far (X. Tang et al., 2020). A synopsis of this review's conclusions indicates that i. there is a lack of CT assessment tools for the upper educational levels, starting from high-school and onwards, ii. the assessment tools should correspond to specific definitions of CT and specific subject-domains of application, iii. they should also be qualitative, iv. undergo a reliability analysis and validation process, v. should support the differentiation of CT from strictly CS-related topics, and vi. should be platform-independent in order to be accessible. Despite a series of limitations, the attempts of measuring CT so far have been numerous (e.g., Ambrósio et al., 2015; Moreno-León et al., 2015; Mühling et al., 2015; Román-González, Moreno-León, et al., 2017; Weintrop et al., 2014; Wiebe et al., 2019), and have paved the way for future progressions in CT research.

Some approaches to measuring CT are tightly dependent on the assessment of projects developed in specific programming environments. Brennan & Resnick (2012b), for example, suggested a CT assessment utilizing the Scratch programming environment for portfolio analysis of projects developed in Scratch, artefact-based interviews, and design scenarios. In the same vein, the *Progression of Early Computational Thinking* model (*PECT*; Seiter & Foreman, 2013) extended the previous approach by integrating evidence variables, design pattern variables, and CT concepts when assessing Scratch projects. Similarly, *Dr. Scratch* (Moreno-León et al., 2015) is a formative CT assessment tool that analyzes automatically CT concepts identified in Scratch projects. There have been quite a few similar approaches dependent on programming environments other than Scratch. The F*airy Performance Assessment* (L. Werner et al., 2012), for example, assesses CT within the Alice programming environment and a real-time assessment tool which, in a like manner, assess CT within activities of the programming environment Agentsheet and Agentcubes (Koh et al., 2014).

Though quite a few examples of CT assessment approaches are linked to specific programming environments, others go beyond both programming environments and programming activities as well. The extensive work of Weintrop et al. (2014), for instance, proposes digital interactive assessment tasks that measure CT within different STEM subjects, like biology, chemistry e.t.c. Likewise, the Organisation for Economic Co-operation and Development (OECD) recently announced that the 2021 PISA (Programme for International Student Assessment) mathematics assessment would incorporate tasks that will assess CT within its mathematics assessment framework (OECD, 1970).

Apart from digital assessments, there have been quite a few psychometric approaches for assessing CT detached from any specific digital environment (e.g., Chen et al., 2017; Mühling et al., 2015). An example of such an approach is the *Commutative Assessment* (Weintrop & Wilensky, 2015), which assesses different CT concepts in two different modalities (i.e., text- and block-based), aiming to provide further insights into the understanding of CT depending on the modality of the tasks.

CT assessments, except from their modality (i.e., digital/non-digital and programming-environment dependent/independent), vary also based on the target group they are designed for. There has been an attempt to assess CT already in pre-school with the *Coding Development* (CODE) *Test 3-6* (Marinus et al., 2018). A very recent assessment tool for the early elementary school is the *Beginners Computational Thinking test* (*BCTt*; Zapata-Cáceres et al., 2020), which is still undergoing validation. A well-validated CT assessment for middle school is the *Computational Thinking test* (*CTt*; Román-González et al., 2017).

According to the population that those tests are addressed to, they utilize accordingly specific graphical designs for their assessment tasks. For example, the *CODE Test 3-6* presents the tasks orally on a life-size mat, and children have to resolve them by programming an educational robot. The *BCTt* employs a kids-friendly graphic environment of mazes on paper that uses colourful animals and shapes for the presentation of the tasks which children have to solve by selecting one of four multiple answers, given as a series of arrows and shapes. Similarly, the *CTt* used a slightly more complex graphic environment to present tasks with more detailed mazes and drawing canvas, which children had to solve by selecting one of four multiple answers, given as a series of visual programming block commands or arrows.

Even though there has been a significant development of CT assessment tools recently, tools that operationalize different definitions of CT, use various modalities and approaches of assessment and are targeting different populations, most of them still luck evaluation and validation with large samples. An exception to this phenomenon is the *CTt*. The *CTt* is the only CT assessment tool so far for which there is empirical evidence on its content validity, criterion validity, convergent validity, predictive validity, and cross-cultural validity (Tsarava et al., 2019).

As Román-González et al. (2019) and Shute et al. (2017) state in their work, there have been a myriad of CT assessment tools developed; however, if we -the research community- want CT to survive and be considered an ability worth to be fostered and developed, we need to define it as a well-established psychological construct. To do so, they suggest to define operational CT models and empirically validate them in order to advance the research on CT assessment and CT in general. Towards this direction, Román-González et al. (2019) classified existing CT assessment tools in seven categories, namely: i. diagnostic tools, ii. summative tools, iii. formative-iterative tools, iv. data-mining tools, v. skill-transfer tools, vi. perceptions-attitudes scales, and vii. vocabulary assessments. They then combined specific CT assessment tools of different categories, and empirically investigated their converged validity. They concluded that the three different types of CT assessment tools they incorporated in their investigation seem to be complementary to each other, and for that reason, they suggested future research on "systems of assessment". These systems would combine tools from different categories instead of using one single CT assessment tool. In that way, CT could be more effectively captured at an overall level.

In conclusion, CT assessment research has made progress during the last years, showing the development of a plethora of assessment tools. Critical for the existence and the development of the CT research field, in general, would be the validation of these tools. Moreover, the elaborated comparative investigation of the existing CT assessment tools would be crucial for the understanding and concrete definition of CT as a cognitive construct of high value for education.

5 Objectives of the Thesis

Addressing the above-described lack of definition, the present thesis seeks to specify CT as a cognitive construct. Therefore, I evaluated associations of CT with other cognitive abilities (e.g., verbal and spatial reasoning, etc.) as well as the contribution of these cognitive abilities to learning CT in a randomized control field trial. Thereby, I gained insights into what CT is as a cognitive construct underlying programming skills as well as how CT may be assessed and fostered efficiently and differentially. A better understanding of CT as a cognitive construct would allow for the design and development of more reliable CT assessment tools and therefore, for evaluated and more effective educational materials for teaching CT.

CT as a term has been widely popularized in the last few decades (Kong & Abelson, 2019; Lockwood & Mooney, 2017; Moreno-Leon et al., 2018), and its importance is notably leading to a considerable development of educational material to foster it (e.g., Grover & Pea, 2013; Lockwood & Mooney, 2017). Nevertheless, these efforts are often evaluated only insufficiently using empirical methods, and the corresponding assessment tools are still in the early stages of development and validation (X. Tang et al., 2020). Against this background, I aim at providing new empirical evidence on CT as a cognitive construct, on the respective assessment of CT, and the effectiveness of specific design approaches for fostering CT. Because CS education has already been part of the elementary school curricula in many countries and is even suggested to be relevant already from the pre-school level, this thesis focuses on an elementary school population.

5.1 Summary of objectives

The overarching aim of this thesis is the cognitive definition of CT, which in the future would allow for more relevant didactical approaches to foster CT, along with the development of appropriate assessment tools for measuring CT. New insights on what CT is in relation to other cognitive abilities will complement the current literature and research on the didactics of CT as crucial 21st-century skillset that needs to be fostered early on in education, either as part of computer science courses or interdisciplinarily in other STEM and non-STEM courses. A better understanding of CT as a cognitive structure will, therefore, facilitate better curricula design, implementation, and evaluation. Thus, it will enhance current and future research on CT assessment tools.

To specify the cognitive correlates of CT, a curriculum was developed for intervention studies. The curriculum design was based on a review of the literature on recent didactic approaches in the CS education community worldwide (Study 1). Accordingly, the unplugged playful parts of the curriculum that were developed exclusively for it were iteratively evaluated to ensure appropriateness regarding the targeted audience‘s age and the study’s purpose (i.e., evaluation with adult participants, Study 2; evaluation with elementary school children, Study 3). Furthermore, an assessment tool for CT in secondary school children (*CTt;* Román-González, Pérez-González, et al., 2017) was adapted and validated to be used as an instrument for assessing CT differentially in comparison to other cognitive skills (Studies 4 and 5). Additionally, the proposed curriculum underwent a multi-level evaluation procedure to evaluate its learning effects as assessed by the newly developed CT assessment instrument, among others (i.e., pilot-phase evaluation, Study 4; effectiveness evaluation, Study 6).

In the following, motivation and objectives of each study will be outlined briefly. The six studies of the current thesis are separated into two sections (i.e., Part II, section 6 and 7) with three studies each. The first section (i.e., 6) focuses on the design and development of the curriculum fostering CT in elementary school children, which was later on used as an intervention for the studies to follow. The second section (i.e., 7) focuses on the specification of CT as a cognitive construct and its assessment. The articles and manuscripts reflecting the research summarized in sections 5.2 and 5.3 are presented in sections 6 and 7, respectively. The individual results of the six studies will be summarized at the beginning of the general discussion (Part III, section 8).

5.2 Curriculum design and development for fostering CT

5.2.1 Study 1: Training computational thinking: Game-based unplugged and plugged-in activities in primary school

In Study 1, a review of the literature was done to identify the current state of curriculum design approaches in elementary school CS education, and upon that to design a new CT curriculum for elementary school children. This curriculum would be offered as an intervention in the following studies (Studies 4 and 6). First, definitions of CT and CT concepts, in particular, were identified from the literature. Then the instructional methods for fostering CT in computer science-related topics were described. Consequently, the conceptualization of a curriculum for fostering CT in 3^{rd} and 4^{th} graders was introduced along with some unplugged/plugged-in playful activities.

The CT concepts identified from the literature were i) sequences, ii) loops, iii) parallelism, iv) events, v) conditionals, vi) operators, and vii) data/variables. These concepts were planned to be introduced in the curriculum using initially playful unplugged activities and then plugged-in ones in order to decrease the cognitive effort of understanding complex concepts and hence increase motivation. These unplugged activities build upon the concept of programmable robots, like Papert's turtle (Papert, 1999), which are non-digital, life-size board games that facilitate embodiment and collaboration during the learning activities. These were proposed to be followed-up by plugged-in activities, which draw on the same CT concepts in a well structured educational programming environment, which facilitates novice programmers by supporting a block-based programming language. All activities addressed topics of different STEM domains and provided a perspective on the usability of CT concepts in real-life problem-solving broader than the context of computer science.

5.2.2 Study 2: Computational thinking through board games: The case of Crabs & Turtles

In Study 2, the unplugged activities of the curriculum conceptualized in Study 1 are described in more detail, and the results of an initial evaluation are presented. The unplugged activities are three life-size board games introduced under the name "Crabs & Turtles: A Series of computational adventures"[8] (in German: Schildkröten & Krabben). They offer a low threshold introduction to CT and coding concepts and are addressed to elementary school children. Moreover, they are designed in life-size embodiment during learning and collaboration during play. In 2018, the games competed at the *12th European Conference on Games Based Learning* and won two prizes, the 1st prize in the category of non-digital games, and the overall 1st prize of the competition (joint). In 2020, they became available as an open educational resource (OER) via the digital repository of OER materials of the University of Tübingen[9].

The study reports the two first phases of the empirical evaluation that the educational games underwent. In these evaluation phases, feasibility and user experience during play were investigated in two different adult samples to ensure the appropriateness of the games before evaluating the target population of elementary school children. First, we examined users' game experience on a sample of university students who provided quantitative and qualitative

[8] https://crabsturtles.iwm-tuebingen.de/

[9] The games are accessible at http://hdl.handle.net/10900.3/OER_MDCKSMXP.

feedback, which was then integrated into the next version of the games. Second, we examined users' gaming experience on a sample of gamification experts and teachers. The results of this study were incorporated into the next version of the games, which was evaluated on a sample of elementary school children (Study 3).

5.2.3 Study 3: Board games for training computational thinking

Study 3 builds upon the results of the previous study, the qualitative feedback of gamification experts and teachers was integrated into a newer version of game instructions, and the users' gaming experience was evaluated on a sample of elementary school children. The iterative process of evaluating the games is presented along with the results of the evaluation on the target population. The results are presented separately for each one of the three games as they were evaluated as separate games. The overall results substantiated the appropriateness of the games as playful activities and were integrated into the curriculum as it was conceptualized in Study 1.

5.3 Cognitive correlates of CT and its assessment

5.3.1 Study 4: Cognitive correlates of computational thinking: Evaluation of a blended unplugged/plugged-in course

In Study 4, a pilot evaluation of the developed CT curriculum is presented as well as an initial investigation of CT and its association with other cognitive abilities. Moreover, the development of a CT assessment tool used in the studies to follow was presented. Course evaluation followed a pre-/post-test design procedure utilizing standardized assessment tools of well-established cognitive abilities (i.e., numerical, verbal, visuospatial) and a CT assessment instrument, which was adapted to fit the study's target population. This instrument resulted as an adaptation of an existing validated assessment tool, which was initially designed for an older group.

The aim of this study was i.) to initially evaluate the feasibility of the course before moving on to the next evaluation phase, which was a randomized control field trial (Study 6) on a larger sample, and ii. to examine the reliability of the assessment tools used for the evaluation and especially the reliability of the adapted CT assessment tool before moving on to a correlational study on a larger sample (Study 5). The results of this study provided first indications of course feasibility and effectiveness, along with a first overview of the relationship of CT with other cognitive abilities.

5.3.2 Study 5: A cognitive approach to defining and assessing computational thinking: An empirical study in primary school

In Study 5, associations of CT with other cognitive abilities in elementary school children are investigated in more depth and relying on a larger sample. Similar research has been done on populations of secondary school, high-school, and university level but not yet on elementary school children. This study aimed at investigating the associations of CT with other cognitive abilities to understand CT as a cognitive construct better and thus enrich its definition and allow for an appropriate assessment approach of CT in this particular age group. The results for associations of CT with other cognitive abilities in this elementary school sample complement the known cognitive correlates of CT across age groups, from elementary school to university level, indicating differences of cognitive interdependencies of CT during age development.

5.3.3 Study 6: Computational thinking training: Implementation and effects on elementary school children's cognitive and computational thinking skills

Study 6 describes the last phase of curriculum evaluation, which is a randomized control field trial on elementary school children—this study aimed at evaluating the effectiveness of the CT curriculum developed as described above. In contrast to the pilot evaluation phase (Study 4), the instructors participating in the study were not the developers of the course material. This approach allowed for conclusions on the effectiveness of the course while controlling for instruction fidelity. The curriculum materials (e.g., lesson plans, short assessments, etc.) along with information about the theoretical and methodological approaches of the course are documented in a detailed course manual of 250 pages (Leifheit et al., 2018). The effects of the course were measured for performance on different cognitive and CT assessments and provided indications on the appropriateness of the course design to foster CT.

Taken together, the six studies as mentioned above, seek to contribute in answering the existing open questions gathered from the research literature of CT. These open questions summarized concern the following: i. a widely accepted *definition* of CT, ii. reliable *assessment tools* for measuring CT abilities, iii. the *cognition* of CT, iv. the appropriate *age* of introducing CT to children, v. the *context* and *modality* of the materials designed to foster CT, vi. the *interdisciplinarity* of the CT concepts, and vii. the *teachers' qualifications* for delivering CT interventions (Angeli et al., 2016; Brackmann et al., 2017; Chiprianov & Gallon, 2016; Lockwood & Mooney, 2017; Sentance & Csizmadia, 2017; Settle et al., 2012; Joke Voogt et al., 2015; P. Wang et al., 2019; Yadav et al., 2014). In Study 1, the literature on CT was reviewed, and accordingly, the conceptualization of a CT curriculum for elementary school students was presented. In Studies 2 and 3, the appropriateness and usability of the game-based materials developed for the CT curriculum was investigated with samples of adults and children. In Study 4, the pilot evaluation of the CT curriculum was conducted, along with a first correlational analysis of CT performance with other cognitive abilities, on a limited sample of elementary school children. Moreover, a CT assessment tool for elementary school students was developed, based on an adaptation of an existing tool for older children. In Study 5, the cognition of CT was investigated by conducting a correlational analysis of CT performance with other cognitive abilities on a sample of almost 200 elementary school students. Finally, in Study 6, the proposed CT curriculum was evaluated for its effectiveness on children's CT abilities in a randomized controlled field trial.

PART II: THEORETICAL & EMPIRICAL STUDIES

6 Computational Thinking: Curriculum Design

In this chapter, the following articles are attached:

- Study 1: **Training computational thinking: Game-based unplugged and plugged-in activities in primary school.**
- Study 2: **Training computational thinking through board games: The case of Crabs & Turtles.**
- Study 3: **Board Games for Training Computational Thinking.**

Training Computational Thinking: Game-based Unplugged and Plugged-in Activities in Primary School

Katerina Tsarava, Korbinian Moeller, Niels Pinkwart, Martin V. Butz, Ulrich Trautwein,

Manuel Ninaus

Abstract

Computational thinking (CT) denotes the idea of developing a generic solution to a problem by decomposing it, identifying relevant variables and patterns, and deriving an algorithmic solution procedure. As a general problem-solving strategy, it has been suggested a fundamental cognitive competence to be acquired in education - comparable to literacy and numeracy. However, integrating CT into general curricula has been challenging. Therefore, the current project aims at developing an extra-curricular training of CT for primary school children. From a literature review, we identified seven concepts central to CT: i) sequencing, ii) loops, iii) parallelism, iv) events, v) conditionals, vi) operators, and vii) data/variables. In our targeted educational training program, we will specifically address these concepts (which are shared concepts between CT and programming/computer science education) in 2-step procedures using corresponding game-based unplugged and plugged-in activities. Playful unplugged activities, such as a treasure hunt board game for the concept of using variables as placeholders for information, shall allow children to get a first grip on CT processes by actively engaging them. In the game, a treasure is to be hunted by completing a series of arithmetic operations, in which players have to handle different variables (e.g., dice faces, scores, etc.). Building on this unplugged activity, a related plugged-in scenario is a programmable simulation of raindrops filling a glass. While raindrop and glass volume are constants, the fill level of the glass may be the variable to manipulate. In both kinds of activities, we aim at clarifying the association between CT-based solving real-life problems and aspects of different STEM disciplines. The series of unplugged and plugged-in activities are integrated into a gamified approach suitable for primary school children, employing badges for mastering specific CT processes to increase students' engagement and give feedback about their learning progress. The instructional design will integrate principles of constructionism, game-based, and project-based learning, such that students will construct knowledge through playing and interacting with interdisciplinary educational scenarios. The course will be empirically evaluated with 3rd and 4th graders in primary schools. Thereby, the idea of evidence-based instruction is pursued to ensure the efficiency and validity of our training.

Keywords: computational thinking, programming, coding, unplugged activities, game-based learning, gamification

1 Introduction

In recent years, there is a growing emphasis on the importance of computer programming or coding skills as 21^{st}-century skills (Wing, 2006, 2010; NRC, 2011). For STEM disciplines, in particular, programming/coding has been argued to be an indispensable instrument for solving complex problems or increasing efficiency through automation (Wing, 2010). Thus, fostering those relevant skills early on in education seems a desirable prerequisite, preparing children for current and future demands of our knowledge societies, spanning from job requirements to leisure time activities. Against this background, the current article proposes an educational course and training for 3^{rd} and 4^{th} graders to foster programming/coding skills. However, in contrast to most similar courses, we take a more cognitive skill-oriented approach, integrating the training of programming/coding skills into the conceptual, theoretical framework of computational thinking (henceforth CT), by employing 2-step procedures using unplugged and plugged-in activities. Moreover, we embedded this in a game-based constructivist pedagogical approach with the aim of introducing CT to young students (by means of coding). CT, as an overarching cognitive skill, is closely related to the different STEM disciplines (e.g., Sanders, 2009). Thus, CT allows for an interdisciplinary perspective on using fundamental coding skills to solve real-world problems. Accordingly, the main contribution of this study will be the development of an integrated framework that fosters coding competence as a practical skill and CT competence as a conceptual cognitive skill.

In the following, we will first elaborate on the close association between programming/coding as a practical and CT as a cognitive skill, before highlighting the relevance of CT for modern educational programs. We then provide a short overview of existing coding and CT trainings, followed by a detailed description of the training we developed to foster coding in 3rd and 4th-graders and a brief conclusion.

1.1 Coding and computational thinking

Computer programming – also referred to as coding – has been coined a crucial 21st-century skill due to the constantly increasing need to keep up with the growing impact of information and communication technologies (henceforth ICT) on human activities. ICT have become prevalent in many facets of everyday life, like production, health, and education, security, job requirements, but also leisure time activities etc. This is reflected in the latest interest of scientific organizations and also governments (e.g., European Schoolnet, European Coding Initiative) all over the world on the establishment of an effective framework for introducing ICT, coding, and CT skills to students already at a young age. Although these three terms share common meanings, they should not be confused as identical. ICT skills refer to general skills related to the use of computer devices and relevant digital content, like software, digital documents, etc. In contrast, coding skills describe the practical ability to write and design software programs as functional computer applications. Finally, CT denotes more general cognitive problem-solving skills based on systematic and computationally-oriented procedures (Balanskat & Engelhardt, 2015). Each of these skills is important not only to become a competent user of ICT but also to meet the needs of our increasingly digitized world. However, while programming/coding is considered a more practical skill, we want to emphasize that CT reflects a broader cognitive concept that is fundamentally critical for becoming computationally literate, besides the fact that at least rudimentary CT is essential for the acquisition of more practical coding skills (Balanskat & Engelhardt 2015; Garcia-Penalvo et al., 2016). At the same time, fostering CT, detached from coding, might result in somewhat subpar and abstract educational scenarios. This fact supports the latest efforts and increased interest in fostering CT as a conceptual cognitive skill that can be applied interdisciplinarily in different domains over the mere training of practical skills, such as coding (e.g., Yadav et al., 2016; see also Figure 1).

Being able to code reflects the "21st-century vision of students who are not just computer users but also computationally literate creators" (https://k12cs.org/). Unsurprisingly, ideas to specifically promote and teach coding abilities already starting in primary school have become increasingly popular (e.g., Balanskat & Engelhardt, 2015; https://code.org/). Central concepts in coding are the generic ideas of sequencing, loops, parallelism, events, conditionals, operators, and data/variables (e.g., Brennan & Resnick, 2012). Interestingly, coding as a practical skill shares these concepts with the psychological construct of CT as a cognitive skill (see Figure 2). Computational Thinking is construed as "the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent" (Cuny et al., 2010). CT denotes the idea of developing a generic solution to a problem by decomposing it, identifying relevant variables and patterns, and deriving an algorithmic solution procedure (e.g., Wing, 2006; Kazimoglu, 2013). In fact, this closely resembles the proceeding in coding. As such, code is usually organized in loops of sequences of defined events that involve specific operations performed on the to-be defined variables. Correspondingly, CT skills specifically draw on processes such as algorithmic thinking, conditional logic, decomposition, abstraction, pattern matching, parallelization, evaluation, and generalization (e.g., Wing, 2010; Briggs, 2014); thereby reflecting cognitive instantiations of concepts central to coding. Importantly, these concepts, as well as their cognitive counterparts in CT, are not to be understood as domain-specific in the sense that they can only be applied to the domain of computer science. Instead, CT should be viewed as a much more general problem-solving strategy, which can be applied to different domains over and beyond computer science (e.g., deductive reasoning). Therefore, CT has been suggested as a fundamental cognitive competence that should be acquired in education – comparable to literacy and numeracy (Yadav *et al.*, 2014).

1.2 Computational thinking in education

As a general problem-solving strategy, the influences of CT are closely related to STEM disciplines (V. Barr & Stephenson, 2011). Furthermore, CT has also begun to influence areas of active study over and beyond STEM, such as algorithmic medicine, computational archaeology, computational economics/finance, digital humanities etc. (Wing, 2010). For this reason, governments and educational institutions all over the world worked on a coherent definition of CT and the integration of CT in the curricula of educational programs of primary, secondary, and higher education over the last decade. For instance, educational institutions in the US revised their undergraduate curriculum in computer science and changed their first course in computer science to cover fundamental principles and processes of CT as a cognitive skill (e.g., Perkovic *et al.*, 2010; Wing, 2010). Moreover, in 2013, the computer science curriculum for universities in the UK was revised by focusing strongly on the promotion of CT as a widely applicable and transferable skill in computer science (Brown et al., 2014). Furthermore, the relevance and importance granted to CT are also reflected by the fact that in 2014 the European Coding Initiative was founded. In collaboration with several European Ministries of Education, members of the European Schoolnet, and the support of major technological enterprises (e.g., Microsoft, Facebook), the initiative aims at promoting a sensible integration and evaluation of coding and CT in the official educational curricula (Balanskat & Engelhardt, 2015).

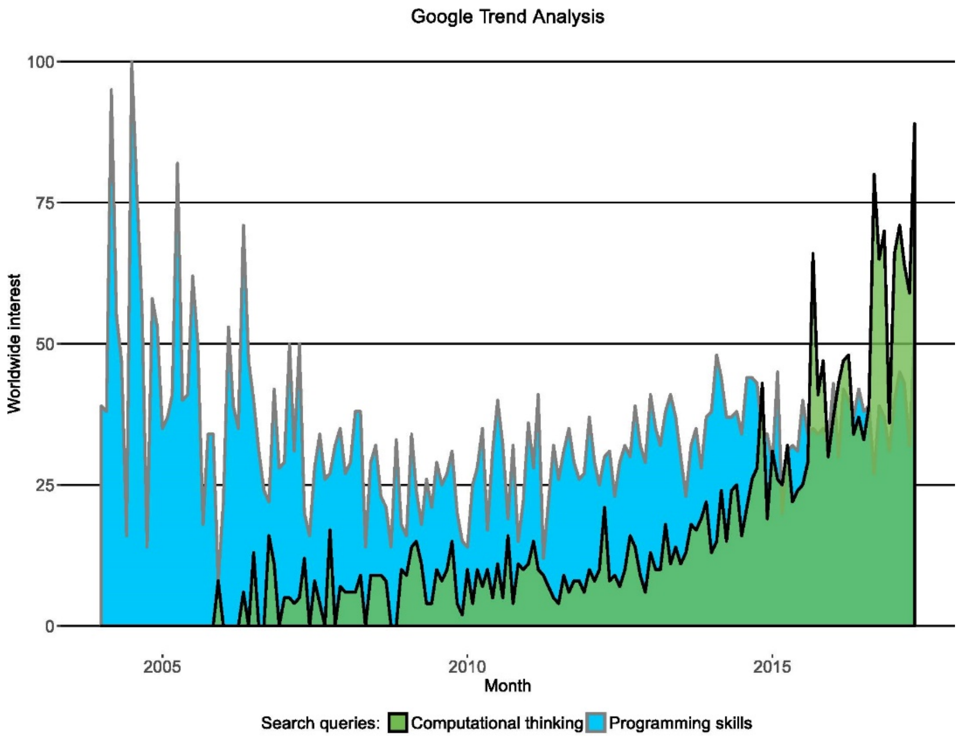

Figure 1: Illustration of Google trends over the course of time, for the search terms "computational thinking" and "programming skills". Worldwide interest (y-axis) reflects the search interest of the corresponding topic relative to the highest point in the chart (https://trends.google.com/trends/). [Accessed 30/04/2017]

This envisaged societal relevance of CT and its wide range of applicability let us decide to develop a training of practical coding skills, integrated into a course on CT applied to various STEM contexts for 3rd and 4th graders. To realize the broad applicability of the training and because of reasons of platform independency, we suggest that coding in young ages should not be based on a specific programming language, as these change rapidly according to market and technology changes. For this reason, we aimed at fostering children's coding skills on the broader and more transferrable level of CT. Moreover, we tried to avoid common concerns on introducing coding already in primary school (Garcia-Peñalvo, 2016) by i) implementing a game-based approach of learning by doing, ii) focusing on cognitive processes of CT and not on practical coding skills related to specific programming languages, iii) using unplugged haptic games and plugged-in low-threshold visual programming environments, and iv) by adopting an overarching gamified framework accompanying the training for maintaining and increasing motivation. Thereby, we build our training on the theory of constructionism following the principles of "learning-by-doing" (e.g., Harel & Papert, 1991), which were established and evaluated in well-known environments for early programming, like the *Logo* programming language, *Scratch*, etc.

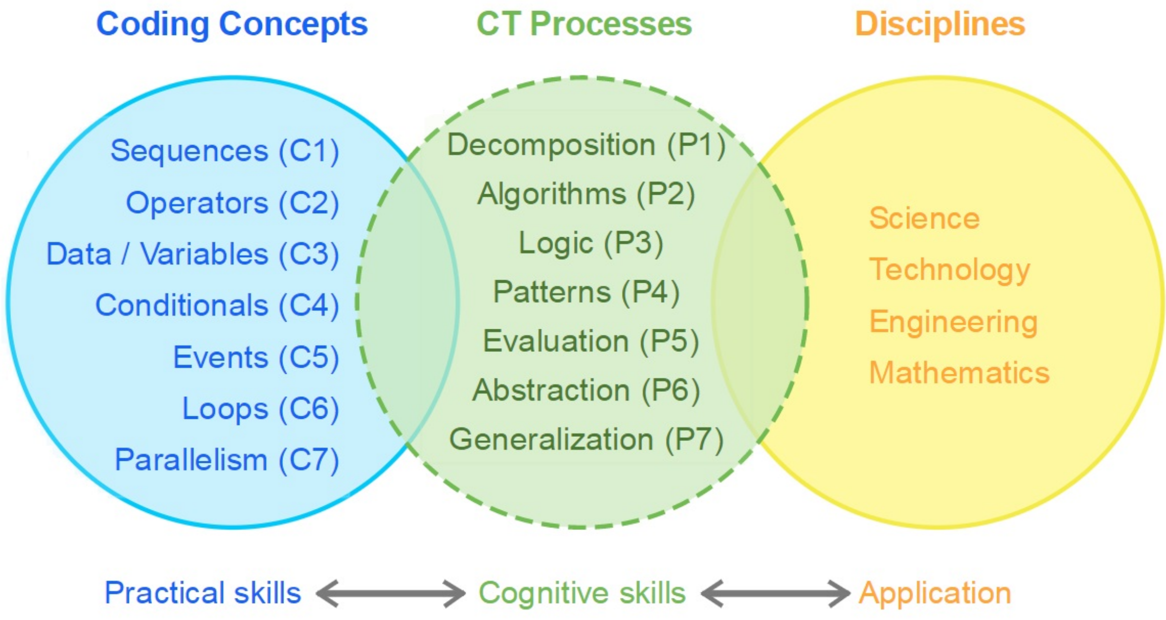

Figure 2: Illustration of association between the practical skills of coding, CT as corresponding cognitive skills and the broad applicability of CT as a general problem-solving strategy to different content domains such as STEM.

2 Course Concept

2.1 Course aim

We specifically designed the course to address CT processes defined and identified as shared with coding. In particular, we considered the concepts of sequences, loops, parallelism, events, conditionals, operators, and data and integrated them into non-programming (i.e., unplugged) and programming (i.e., plugged-in) activities. The instructional design of our training is based on introducing each of the CT processes in a multimodal way, using unplugged and plugged-in activities, and demonstrating their applicability within different STEM-related contexts. The general idea of the whole course follows the theme: "play-modify-create". Students are introduced to CT processes through playful unplugged activities. Subsequently, they are asked to modify elements within existing plugged-in activities before they finally have to create their own usable designs.

In the following description of the course concept, we first describe the actual lesson content and activities and their aim. Subsequently, we elaborate on how the employed activities allow for a broad applicability of CT by relating the activities to different STEM contexts. Moreover, we outline how the combination of unplugged and plugged-in activities allows for an integrated constructivist approach to convey the respective content. Finally, we briefly describe how we use a gamification framework to incorporate lessons on differing content conveyed in different modes into a coherent and overarching course design.

2.2 Course outline

The course is structured as a series of eight lessons of 90 minutes each (see also Figure 3). During these lessons, CT processes are introduced gradually, beginning from more unplugged haptic, practical, and experiential activities, moving on to plugged-in more abstract and demanding ones. During the lessons, students create their own applications with MIT *AppInventor,* which they can reuse on their own devices. Teacher's guidance is gradually decreased towards students' gradual independence of learning and creating. The specific lesson plan is as follows:

2.2.1 Lesson 1

Description: Students are first introduced to the gamified assessment framework (see below). Moreover, they get acquainted with unplugged concepts and tangibles and are introduced to the idea of computing without a computer. The first activity is an unplugged life-size board game with turtles. The game shares ideas with the concepts of the educational programming language *Logo* and is inspired by the commercial board game *Robot Turtles* (Shapiro, 2013). In this treasure hunting game, small groups of students have to manipulate turtle pawns, which move by following specific commands written on game cards. Players need to edit and combine command-cards and make strategic decisions to create effective sequences, which allow them to lead their pawns to the place where a treasure can be found. The aim of the game is the fast and efficient collection of treasure items.

Aim: The main purpose of this first activity is the playful introduction to CT processes, such as logical and algorithmic thinking, as well as pattern recognition through the use of common coding mediums, like sequences and loops.

2.2.2 Lesson 2

Description: This lesson encompasses playing within unplugged activities and recognizing CT processes in STEM disciplines. The second activity also employs an unplugged treasure hunt like a board game and utilizes math problems as progression stages. Specifically, this multiplayer board game is a competitive scenario, where groups of players have to find their way through a maze of various difficulty levels. Each challenge includes equations containing variables and placeholder images for constants (e.g., a blue crystal reflecting a value of 4). Small groups of players have to solve the respective arithmetic equations in alternating order and find the best strategy to progress on their way to the centre of the maze. Conditions set by the game's board (maze) provide obstacles to obstruct the most direct way of reaching the centre.

Aim: This activity aims at introducing conditionals, operators, variables, and constants, as well as previously presented coding concepts (i.e., sequences and loops) to foster CT skills of logic, algorithmic thinking, and evaluation.

2.2.3 Lesson 3

Description: Students play within unplugged activities and concepts, which are then gradually transferred to the plugged-in environment of *AppInventor*, reusing established concepts from previous lessons. In this blended activity, students have the opportunity to observe how unplugged coding and CT processes, like, for example, events and parallelism, are applied and how they function within the plugged-in programming environment through simple precoded scripts and scenarios. For instance, in a science simulation about rain (event) drops (variable) increasing the fill level (variable) of a glass (constant), students need to recognize the coding concepts previously introduced unplugged and understand how they are depicted and used in the plugged-in environment. Students should be able to recognize, use, and modify coding concepts in these pre-built *AppInventor* applications.

Aim: In this lesson, students should comprehend the interconnections between coding concepts and the newly introduced CT abstract processes of decomposition and generalization.

2.2.4 Lesson 4

Description: After recapitulating the coding concepts and CT processes already introduced, Lesson 4 requires students to brainstorm real-life scenarios and applications of these concepts, to highlight the importance of CT processes in everyday life and STEM disciplines in particular. Following this, students are introduced to the *AppInventor* software through multiple interactive tutorials as well as editing and playing simple game applications. Pre-developed simple games in the MIT *AppInventor* environment are used as demonstrators and allow students to manipulate code elements, e.g., building blocks/variables, in order to grasp the effects of changes in a running system.

Aim: The activities of this lesson are intended to support students' familiarity and understanding of the environment and how visual coding blocks can replicate coding concepts already identified in the previous lessons.

2.2.5 Lesson 5 & 6

Description: In Lessons 5 and 6, students are guided through the creation of a simple app using scenarios in STEM contexts. A simple calculator is developed by first explaining and understanding its usability and later on designing and programming it in *AppInventor*. Afterwards, students are asked to create other and more advanced apps in other STEM disciplines, for instance, science simulations. By providing pre-built *AppInventor* assets to students, we can facilitate work and guide the learning experience even in complex projects. Different projects are assigned randomly to small groups of students. For instance, the creation of a simple pool billiard physics app, to understand and visualize kinetic energy/momentum conservation of colliding balls. Other projects require, for example, the creation of apps that simulate a magnetic field and the forces operating in it, and the creation of the four seasons, or how the water cycle works, etc. After completing their respective app, all the student groups have to interact and test the creations of their peers.

Aim: Both activities aim at fostering students' coding independence through fading out teacher guidance. The CT processes fostered by these activities are the process of evaluation and abstraction. Moreover, using and developing simulated real-life STEM contexts should increase the awareness of the necessity of coding skills in order to solve problems in different STEM disciplines.

2.2.6 Lesson 7

Description: During the 7th lesson, students are asked to brainstorm simple game ideas. Once they decided on one of the designs, students can create their own game. Of course, they need to create rather simple games (e.g., dice, memory game, mini-golf, etc.) to keep it feasible.

The instructor is crucial in this part of the lesson as he/she has to assist in deciding on a realizable game by taking into consideration its basic mechanisms. After deciding on a game, students collaboratively create the game application. Importantly, in such complex projects, students have to apply all their previously learned coding and CT skills by understanding, analyzing, designing, implementing, and evaluating their own game app. Created games can also be shared among group members and peer-evaluated by fellow groups of students.

Aim: The aim of this activity is the engagement of students with more complex activities of problem-solving and procedural thinking, by creating and evaluating designs of their own.

2.2.7 Lesson 8

Description: In the last lesson, students have to create their own applications. They are asked to create an application dedicated to one of the STEM disciplines already presented and adapt or extend existing programs. They are urged to do so by reusing parts of code created in previous lessons to facilitate the working process, but should also integrate new mechanics or functionalities, respectively, into the application; for instance, scripts of app interface functionalities, such as interactive screen components, random number generators, etc.

Aim: During this activity, students also have to follow the procedure of analyzing the demands and requirements to design an effective structure for their app. The evaluation procedure relies on sharing and peer-reviewing, as beta testers will test the apps of fellow student groups, repeating and fostering the obtainment of all the previously identified CT processes.

2.3 Unplugged and plugged-in activities

Contemporary board games have proven to represent an informal and interactional context in which computational thinking has to be applied. For instance, *Pandemic* (Leacock, 2008) and *RaBit EscApe* (Apostolellis et al., 2014) are two strategic board games, in which computational thinking was embedded in collaborative play. Considering this evidence, unplugged activities employed in the present course are realized as life-size board games, in which students play collaboratively around a floor-board by strategically solving problems and manipulating their pawns accordingly in space. Their active engagement in those unplugged games should raise their motivation for participation and learning (see Echeverría et al., 2011; for an overview), as well as allowing for an embodied experience of basic coding concepts and

CT processes (cf. Barsalou, 2008 for embodied cognition), supporting conceptual abstractions in a natural manner (Butz, 2016). Moreover, the game-based approach of the employed plugged-in activities does not only aim at engaging students into the learning activities but should also enhance the training and development of students' symbolic thinking through multimodal representations (Plass et al., 2015) and simplifications of complex computer-related concepts (e.g., the concept of variables and constants described above, represented by the game rules as objects of predetermined value).

For plugged-in activities of the course, we selected the MIT *AppInventor* software in its browser-based version. MIT *AppInventor* offers a novice-oriented introduction to programming and app creation that transforms the complex language of text-based coding into visual drag-and-drop building blocks. The low-threshold graphical interface allows even an inexperienced novice to create a basic, fully functional app within an hour or less. *AppInventor* allows the development of applications for Android-run devices, using a web browser and either a connected smartphone/tablet or emulator. This allows for taking home self-generated apps as a trophy after the learning activity. We consider this software an advanced alternative to Scratch visual-programming language, as it allows the creation and distribution of a standalone application.

The design of the course embeds the training of CT skills in a multimodal procedure. Coding concepts and associated CT processes are first introduced in a playful and embodied way (unplugged activities), before they are reconsidered in programming context (plugged-in activity), which also implies their application in a STEM discipline. This aims at highlighting the relevance of coding concepts and CT processes not only for digital contexts but also real-life problems in general and STEM contexts in particular. For instance, in Lesson 2, students play a math-based treasure hunting game. Following the rules of the game, players have to devise effective sequences of commands by combining constants, variables, and operators correctly. They have the opportunity to make their sequence even more successful by recognizing patterns of moves, which may be folded and operated by loops. Those unplugged game rules reflect fundamental and applicable coding concepts, which can easily be applied and transferred to any programming language or complex problems in STEM. Accordingly, the aforementioned activity is integrated into a plugged-in task in Lesson 3, where several of these coding concepts are integrated into short, simple pre-coded scripts. As an advanced task, students are then asked to modify those scripts experimentally to observe and experience the immediate consequences of their changes (picking up on the idea of live coding, e.g., Paxton, 2002).

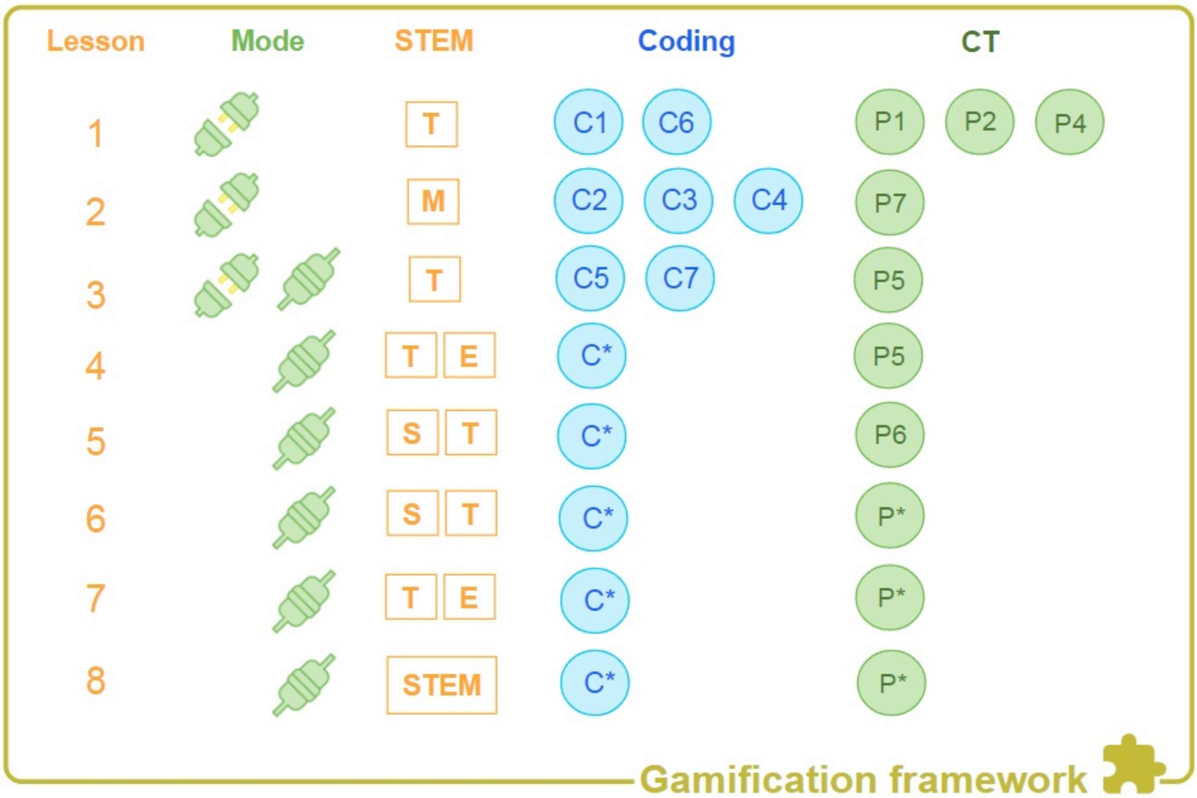

Figure 3: Illustration of the course design taking into consideration the factors of mode (i.e., unplugged/plugged-in), coding concepts (C1-Sequences to C7-Parallelism, see Figure 2), CT processes (P1-Decomposition to P7-Generalization, see Figure 2) and the gamification framework. C*/P*: all concepts (C1-7) and processes (P1-7).

2.4 Gamification and assessment framework

By employing digital games as a learning medium and providing an overarching gamification framework for the course, we aim at increasing the motivation and enjoyment of students. In fact, as a medium for learning, games provide promising possibilities to motivate and engage students in learning (e.g., Chen et al., 2012). Importantly, even simple game-like extrinsic motivators, such as score points and badges, can increase enjoyment and performance (e.g., Ninaus et al., 2015; for a review, see Hamari et al., 2014). In the current CT course, we use a gamified assessment framework, which is based on the assessment framework created by Dorling and Walker (2014) for the effective evaluation of the UK computer science and CT curriculum (see also Moreno-León et al., 2015). As such, we apply a gamified award system, awarding badges for the successful acquisition of coding concepts, core CT processes, STEM specialization domain (e.g., leaderboard for Maths, Science, etc.), creativity, and social skills (e.g., cooperation within or between group, etc.). For instance, students receive stickers for each attended session to put down in their individual course membership card, or, after successfully creating a science simulation, students are awarded a science-badge.

3 Future Studies and Conclusion

The current course is planned to be part of the Hector-Core-Course program of the Hector-Children Academies[10] in Germany, which provides extra-curricular enrichment programs. Therefore, the course will undergo a rigorous three-stage evaluation process. Phase 1 will include piloting and testing the course concept. For this reason, multiple rounds of discussion with experts on the content as well as educators, will take place. This phase also includes an initial evaluation of the effectiveness of the course in a small-scale intervention study at about 3-6 Hector-Children Academies to acquire first empirical data on training gains and the feasibility of the course design. In phase 2, feedback and experiences generated in phase 1 might result in modifications of the course. Following this, another empirical evaluation of the effectiveness of the course at about 10 Hector-Children Academies will be run using a pre-post-test control group design. Importantly, in this phase, we will also evaluate the training of instructors as well as the training material itself. Finally in phase 3, implementation and effectiveness of the course will be evaluated in a randomized controlled field trial involving at least 20 Hector-Children Academies. Evaluation in phases 2 and 3 will also aim at assessing possible transfer effects of the training by employing standardized psychological tests in order to examine whether CT training affects other related cognitive skills, such as reasoning or general problem-solving skills. The primary objective at this stage of the evaluation is to whether the course yields any overall effect on computational thinking. Given that positive effects are observable in all three evaluation phases, the current course will be certified as a Hector-Core-Course to be offered to all Hector-Children Academies. Moreover, in order to better understand the underlying mechanisms on how the various elements of the course influence the overall efficiency, design research methods will be applied. After the design, development, and evaluation phases, we expect to deliver hands-on un-plugged games and their related plugged-in activities built in *AppInventor*, along with instructional materials for future teachers of the course.

---

[10] www.hector-kinderakademie.de

The design and development of the current course is based on the most recent literature on educational practices for coding and CT introduction into official curriculums and latest educational practice in STEM. We integrated elements of game-based learning and gamification methods aiming at engaging and motivating students, while specifically addressing STEM context to reflect the broad applicability of CT. Importantly, this course does not aim at being a core programming course. Using a more general and cognitive perspective on programming and coding we aim at fostering the underlying cognitive concept of CT, which might have broader beneficial effects than instructing a single programming language alone. Consequently, the course is not only aiming at improving practical programming skills but fundamental cognitive skills relevant for the 21st century.

References

Apostolellis, P., Stewart, M., Frisina, C. and Kafura, D. (2014) 'RaBit EscAPE: A Board Game for Computational Thinking', *Proceedings of the 2014 conference on Interaction design and children - IDC '14*, pp. 349–352.

Balanskat, A. and Engelhardt, K. (2015) 'Computing our future. Computer programming and coding Priorities, school curricula and initiatives across Europe', Brussels: European Schoolnet, pp. 4-7.

Barr, V. and 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), pp. 48–54.

Barsalou, L. W. (2008) 'Grounded cognition', *Annual review of psychology*, 59, pp. 617–645.

Brennan, K. and Resnick, M. (2012) 'New frameworks for studying and assessing the development of computational thinking', *Annual American Educational Research Association Meeting, Vancouver, BC, Canada*, pp. 1–25.

Brown, N. C. C., Sentance, S. U. E., Crick, T. O. M. and Humphreys, S. (2014) 'Restart : The Resurgence of Computer Science in UK Schools', 14(2).

Briggs, J. (2014). Computational Thinker: Concepts & Approaches. CAS Barefoot. Retrieved from http://barefootcas.org.uk/wp-content/uploads/2016/08/Barefoot-Computational-Thinking-Poster.pdf .

Butz, M. V. (2016) 'Toward a unified sub-symbolic computational theory of cognition', *Frontiers in Psychology*, 7, pp. 1–19.

Chen, Z. H., Liao, C. C. Y., Cheng, H. N. H., Yeh, C. Y. C. and Chan, T. W. (2012) 'Influence of game quests on pupils' enjoyment and goal-pursuing in math learning', Educational Technology and Society, 15(2), pp. 317–327.

Cuny, J., Snyder L. and Wing J. M. (2010) "Demystifying Computational Thinking for Non-Computer Scientists" work in progress.

Echeverría, A., García-Campo, C., Nussbaum, M., Gil, F., Villalta, M., Améstica, M. and Echeverría, S. (2011) 'A framework for the design and integration of collaborative classroom games', Computers & Education. 57(1), pp. 1127–1136.

García-Peñalvo, F., Reimann, D., Tuul, M., Rees, A. and Jormanainen, I. (2016). TACCLE 3, O5: An overview of the most relevant literature on coding and computational thinking with emphasis on the relevant issues for teachers. pp. 3-8.

Google Inc. and MIT Media Lab (2010) 'AppInventor 2'. Available from http://ai2.appinventor.mit.edu.

Harel, I. and Papert, S. (1991) 'Constructionism', Ablex: Norwood, NJ.

Hamari, J., Koivisto, J. and Sarsa, H. (2014) 'Does gamification work? - A literature review of empirical studies on gamification', Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 3025–3034.

Kazimoglu, C. (2013) 'Empirical evidence that proves a serious game is an educationally effective tool for learning computer programming constructs at the computational thinking level'.

Leacock, M. (2008) 'Pandemic' [Board game], Z-Man Games: Mahopac, NY.

Moreno-León, J., Robles, G. and Román-González, M. (2015) 'Dr. Scratch: Automatic Analysis of Scratch Projects to Assess and Foster Computational Thinking', *RED. Revista de Educación a Distancia*, 15(46), pp. 1–23.

NRC (2011) 'Committee for the Workshops on Computational Thinking: Report of a workshop on the scope and nature of computational thinking', Washington, DC: National Academies Press.

Ninaus, M., Pereira, G., Stefitz, R., Prada, R., Paiva, A. and Wood, G. (2015) 'Game elements improve performance in a working memory training task', *International Journal of Serious Games*, 2(1), pp. 3–16.

Paxton, J. (2002) 'Live programming as a lecture technique', *Journal of Computing Sciences in Colleges*, 18(2), pp. 51–56.

Perkovic, L., Settle, A., Hwang, S. and Jones, J. (2010) 'A Framework for Computational Thinking across the Curriculum', *Proceedings ITiCS '10*, pp. 123–127.

Plass, J.L., Homer, B.D., Kinzer, C.K., Plass, J.L., Homer, B.D., Kinzer, C.K., Plass, J.L., Homer, B.D. and Kinzer, C.K. (2016) 'Foundations of Game-Based Learning', *Educational Psychologist,* 50(4), pp. 258–283.

Sanders, M. (2009) 'STEM, STEM Education, STEMAnia', *Education*, 68(4), pp. 20–27.

Shapiro, D. (2013) 'Robot Turtles', Thinkfun: Seattle, WA.

Wing, J. M. (2006) 'Computational Thinking', *Communications of the Association for Computing Machinery (ACM)*, 49(3), pp. 33–35.

Wing, J. M. (2010) 'Computational Thinking: What and Why?', T*he Link - The Magazine of the Carnegie Mellon University School of Computer Science*, pp. 1–6.

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S. and Korb, J. T. (2014) 'Computational Thinking in Elementary and Secondary Teacher Education', *ACM Transactions on Computing Education*, 14(1), pp. 1–16.

Yadav, A., Hong, H. and Stephenson, C. (2016) 'Computational Thinking for All: Pedagogical Approaches to Embedding 21st Century Problem Solving in K-12 Classrooms', TechTrends, 60(6), pp.565-568.

Computational Thinking through Board Games: The Case of Crabs & Turtles

Katerina Tsarava, Korbinian Moeller, Manuel Ninaus

Abstract

As a cognitive ability computational thinking describes a specific way of algorithmic reasoning building on concepts and processes derived from computer programming/coding. Recently, computational thinking was argued to be a fundamental and educationally relevant 21st-century skill that should be fostered already in childhood. Accordingly, we developed three life-size board games – Crabs & Turtles: A Series of Computational Adventures – aimed at providing an unplugged and low-threshold introduction to computational thinking. In particular, the games aimed at introducing basic coding concepts and computational thinking processes to 8 to 9-year-old primary school children. In the current study, we first describe the design of the games in detail to explicate the development process and allow for reproducibility. We then report on a first empirical evaluation of feasibility and user experience of our educational board games in a two-phase approach. We conducted quantitative analyses of player experience and qualitative feedback of adult student participants (phase 1) and a sample of gamification experts and teachers (phase 2). We examined users' game experience with an adult population to ensure the game's appropriateness. Results indicated an overall positive game experience for all three games. Future studies would be desirable, which should evaluate player experience and learning outcomes in the primary target population of children.

Keywords: educational board games, computational thinking, coding, embodied cognition

1 Introduction

Computational Thinking (CT) denotes the idea of developing a generic solution to a problem by decomposing it, identifying relevant variables and patterns, and deriving an algorithmic solution procedure (Wing, 2006a). As such, CT represents a cognitive ability to apply fundamental concepts and reasoning that derive from computer science in general and computer programming/coding in particular to different other domains, including real-life activities (Wang, 2015). Accordingly, CT is considered a fundamental ability for everyone and not just for computer scientists (Wing, 2006a). The psychological construct of CT as a cognitive ability shares common concepts with coding as a practical skill. Central concepts in coding are the generic ideas of sequencing, loops, parallelism, events, conditionals, operators, and data/variables (Brennan & Resnick, 2012a). Correspondingly, CT abilities specifically draw on processes such as algorithmic thinking, conditional logic, decomposition, abstraction, pattern matching, parallelization, evaluation, and generalization (Astrachan & Briggs, 2012; Wing, 2010); thereby reflecting cognitive instantiations of concepts central to coding.

Importantly, these concepts, as well as their cognitive counterparts in CT, are not to be understood as domain-specific in the sense that they can only be applied to the domain of computer science. Instead, CT should be viewed as a more general problem-solving strategy, which can be applied to different domains over and beyond computer science. Therefore, CT has been suggested to be a fundamental cognitive ability that should be acquired in education – comparable to literacy and numeracy (Yadav et al., 2014).

This broad applicability of CT abilities has lately led to several adaptations and reformations of educational programs (e.g., in Finland where coding was introduced as a subject recently; see Brown et al., 2014; Tuomi et al., 2018). Governments and educational institutions all over the world have been working on a coherent definition of CT and its integration in curricula of educational programs of primary, secondary, and higher education (e.g., Brown et al., 2014; *Code.Org*.; *European Coding Initiative*; *European School Network*; *National Science Foundation*). This envisaged societal relevance of CT and its wide range of applicability inspired us to develop a training of practical coding skills integrated into a course on CT applied to various STEM contexts for 3rd and 4th graders (for the overall course program, see Tsarava et al., 2017). This approach aims at highlighting the relevance of coding concepts and CT not only for digital contexts but also real-life problems in general and STEM contexts in particular, thereby increasing students interest in improving their CT skills. Moreover, taking into consideration common concerns on introducing coding already in primary school (Garcia-Peñalvo, 2016), we aimed at implementing a game-based approach of learning by doing, focusing on central concepts of CT and not on practical coding skills related to specific programming languages. To do so, we developed and employed, among others, unplugged life-size board games *Crabs & Turtles: A Series of Computational Adventures* (henceforth referred to as *Crabs & Turtles*). A first empirical evaluation of these will be described in the current article. Games or game-based applications are an increasingly important approach in cognitive training, learning, and educational interventions because of their ability to keep learners motivated to play and to interact with the application or learning environment, respectively (Boyle et al., 2016a; Plass et al., 2016). Recent research even indicated that game-based learning might be more effective in terms of learning and retention than conventional instruction methods (Wouters et al., 2013).

Our game design relied on Piaget's theory of constructivism (Papert, 1999) and was further inspired by Papert's integrated constructionism approach (Kafai & Burke, 2015; Papert, 1999). In addition, we were inspired by the successful implementation of the haptic Logo-Turtle (Papert, 1999; Papert & Solomon, 1971), which led to the Logo visual programming language. As regards content, we considered the central concepts of coding as identified by (Brennan & Resnick, 2012a). After years of development and evaluation using the educational software Scratch, they identified seven overarching computational concepts, applicable to other programming and non-programming contexts but also generalizing beyond them: (i) Sequences, (ii) Loops, (iii) Parallelism, (iv) Events, (v) Conditionals, (vi) Operators, and (vii) Data. We integrated 6 of those concepts into our game content design and aimed at training children through un-plugged playing activities in a board game (Table 1). Here, we describe the development, design, and results of initial user tests of three games – all addressing different CT concepts – which are subsumed under the game series *Crabs & Turtles.*

While there are a number of games aiming at training CT related abilities, most of them are digital (e.g., *Program your Robot,* Kazimoglu, 2013), whereas only a few allow for non-digital haptic (e.g., *Robot Turtles,* Dan, 2013; *Ricochet Robots,* Randolph, 1999; *Pandemic,* Leacock, 2012), and thus embodied or blended approaches (e.g., *Osmo Coding Family*). Moreover, these games can be further distinguished on whether they are a commercial (e.g., *Qwirkle,* McKinley, 2006) or research and experimental project productions (e.g., *Dragon Architect,* Bauer et al., 2015; *Rabbit Escape,* Apostolellis et al., 2014). All of them differ in their target audience, holistic perspective, and mode. *Program your Robot,* for example, is a web-based environment aiming at introducing computer programming concepts and various CT skills, such as problem-solving, algorithm building, debugging, etc. In the game, players have to manipulate non-verbal commands by dragging and dropping them to program their robot to collect or avoid items. *Dragon Architect* is another web-browser game based on the *Blockly* (Fraser, 2015) programming environment. It introduces concepts through puzzles that require a command solution, which gradually becomes more difficult. There are quite a few games like those aiming at supporting CT skills with promising results so far (Berland & Lee, 2011; Kazimoglu et al., 2012). However, most of them lack qualitative and/or quantitative evaluation of their training effects.

Crabs & Turtles shares common ideas with concepts of the educational Logo-Turtle and logo-inspired games and gamified educational activities. Importantly, the development process of the game was driven by own previous research and piloting. For instance, in 2016, we created a life-size board game called *Turtle Steps* (Tsarava, 2016), which can be considered the initial archetype educational intervention of *Crabs & Turtles*. The game aimed at an embodied training of simple computational concepts with direct transferability to an educational Python editor environment, in which children were able to program in a native translation (Greek) of the actual Python programming language. After multiple pilot sessions with *Turtle Steps*, we derived conceptual ideas for the first game-based learning activities of *Crabs & Turtles*. Note that we intentionally designed *Crabs & Turtles* to be independent of any specific programming environment or language. The games' main target group are primary and secondary school students (8-12 years old) with no prior programming knowledge. It is, however, also suitable for older students and adults with no programming experience. The life-size dimensions of the game allow playability within the classroom or open-air spaces, such as a schoolyard. We chose the life-size game design to encourage active engagement and participation and thus to increase children's motivation for active learning (for an overview, see Echeverría et al., 2011), on the one hand, and to enhance learning outcomes by an embodied experience of basic coding concepts and CT processes (cf. Barsalou, 2008 for the concept of embodied cognition) supporting conceptual abstractions (Butz, 2016), on the other hand. The chosen un-plugged mode takes into account common concerns on introducing coding to primary school children (e.g., Grover & Pea, 2013; Pea & Kurland, 1984) and offers a smooth and children friendly transition to digital, more complex educational programming environments. Moreover, we feel that using a non-digital mode is crucial because it fosters the experience that possible applications of coding concepts and CT processes are not restricted to digital contexts but also generalize to real-life conditions (Tsarava et al., 2017). Although the game can be used as a standalone game intervention, it is intended to be part of a structured course curriculum (Tsarava et al., 2017), which builds upon skills acquired within the game. To build our game, we followed an iterative user-centred development process (Fullerton, 2008). In particular, first design ideas of the game content were tested with a custom-made life-size game as a pilot educational intervention with primary school children (Tsarava, 2016). Next, an early prototype was developed and tested in terms of usability of the materials needed to play the game (e.g., printed wooden floor tiles vs linoleum canvas). During a 2 hour workshop with children, qualitative feedback was gathered and used to further improve the overall design. Finally, we examined users' game experience quantitatively with an adult population to ensure the game's appropriateness. After providing a detailed description of *Crabs & Turtles*, the presented article reports the results of two evaluation studies with adults. Game experience was evaluated in two phases of playtesting with (i) a general audience of postgraduate students and went on to (ii) a more specialized group of gamification experts and teachers. In the following, we will first describe the design of the latest version of the game before reporting the results of the user experience studies afterwards.

2 Game Description

Crabs & Turtles consists of three different games: i. *The Treasure Hunt*, ii. *The Race*, and iii. *Patterns*. Currently, all three games are available in English, German *("Schildkröten & Krabben"*), and Greek *("Χελώνες & Κάβουρες"*). It is primarily designed for children at the primary school level. The teachers or educators play an important role in each of these games and are in close contact with their pupils by acting as game masters. The games aim at training cognitive processes related to CT, such as algorithmic thinking, abstraction, pattern recognition, and decomposition (see Table 1). These processes can be either applied to specific coding skills (i.e., sequences, loops, conditionals, patterns, and events) or to mathematical skills (i.e., dealing with angular degrees in spatial orientation, addition, and multiplication) as well as skills relevant to both coding and mathematics (i.e., operators, variables, constants, and values). Our game design can be described within the framework for educational game design as proposed by Roungas & Dalpiaz (2016). The game design elements of goals, game mechanics, and challenges were carefully selected and adapted when necessary, as was the element of feedback for each game decision given by a teacher serving as the game master in all the three games. Moreover, elements and design decisions related to educational games like curriculum, readiness for learning, stimuli, and rewards were cautiously selected.

Below we describe the game design of the three different games in detail and specify which cognitive processes and learning objectives, respectively, are primarily addressed in each of them.

Figure 1. The Treasure Hunt game (1. Sequence board, 2. Game starting points, 3. Game pieces, 4. Water grid, 5. Stone grid, 6. Grass grid, 7. Treasure collection location).

2.1 The Treasure Hunt

The Treasure Hunt is the first game of *Crabs & Turtles*. In this game, users have to manipulate coloured game pieces representing turtles and crabs to figure out the most efficient way to collect treasures placed on the grid squares of the game board (see Figure 1). To move a crab or turtle, users need to create effective sequences of commands on a sequence board (see Figure 1 & 3), representing specific coding concepts. For instance, users have to build sequences of steps, turns, and loops to move their game piece towards treasures fast and efficiently. As such, coding concepts and, to a lesser degree, mathematic abilities trained by this game are sequence building, value understanding, dealing with angular degrees and spatial orientation, loop creation, as well as conditional decisions.

Importantly, there are some restrictions with regard to the game board and game pieces (turtle vs crab) that affect the players' strategy. For instance, turtles can move only on grid squares indicating stone and grass grounds, whereas crabs can only move on grid squares representing stone and water ground (see Figure 2). Additionally, turtles can move only forward and backwards, contrary to crabs that can move only sideways to either the left or right. However, under specific conditions, for instance, when an Event card from the pile, indicating *Walk forward/backward*, is picked up from a crab team, it can be used as it would be used by a turtle. Those cards are considered bonus cards for crabs, and correspondingly *Walk left/right* are bonus cards for turtles because they allow movement in more directions than usual.

Table 1. Coding skills & CT processes as aimed game trained skills.

	Coding Concepts						CT Processes				
	Sequences	Operators	Constants / Variables	Conditionals	Events	Loops	Decomposition	Algorithms	Patterns	Evaluation	Abstraction
The Treasure Hunt	x					x	x	x		x	
The Race		x	x	x	x						x
Patterns									x		

2.1.1 Learning objectives

The main learning objectives of this game are the general introduction to algorithmic thinking, the use of commands in specific and sequential order, and the consideration of restrictions by possible conditions when forming a strategic solution to a problem. After playing the game, we expect users to have acquired an understanding of what simple algorithms are and how they are formed as sequences of several commands that serve a specific strategical purpose.

Besides that, we expect participants to be able to consider specific restrictions when making their decisions and recognize small repeatable patterns that can be folded into a loop.

Table 2. The Treasure Hunt game cards

Motion command cards		Event cards
Crabs	Turtles	
turn ↷ ___ °	turn ↷ ___ °	walk forward 2
turn ↶ ___ °	turn ↶ ___ °	walk backward 2
walk left ___	walk forward ___	walk left 2
walk right ___	walk backward ___	walk right 2
repeat 2	repeat 2	turn ↷ 90°
		turn ↶ 90°

2.1.2 Game play and rules

The game is played in teams of two. Each team possesses one game piece (crab or turtle), a sequence board, five re-writable motion command cards (see Table 2), a marker, and a sponge (see Figure 3). Each turn, teams draw one more Event card from the pile (see Table 2), which they can either use in building their sequence on the sequence board or return it at the end of their turn. The goal of the game is to collect a specific number of food-treasures (e.g., three), which are spread across the game-board, as fast as possible. Each team has to collect three different food items from three different treasure points on the board grid. Turtles collect magenta-coloured items, while crabs collect green ones (Figure 3). To approach the treasure on the grid of the game board, teams have to structure their command cards on their sequence board and, at the same time, consider the respective restrictions for crabs and turtles (i.e., crabs can move only on stone and water grid squares, while turtles can move only on grass and stone grid squares). Both types of game pieces can step on all treasure locations (see Figure 2). At each turn, teams have in total 6 cards from which they can use a maximum of 5 in order to build their sequence. When executing a sequence of commands, users are rewarded skill badges related to their achievements. These are collected on the back of the game pieces (see Figure 3 and Scoring). The first team to fulfil the condition to win the game (e.g., collect three pieces of food) has to wait for the round to finish. In case there is another team to achieve the winning condition within the same round, the winner is determined by the number of badges collected.

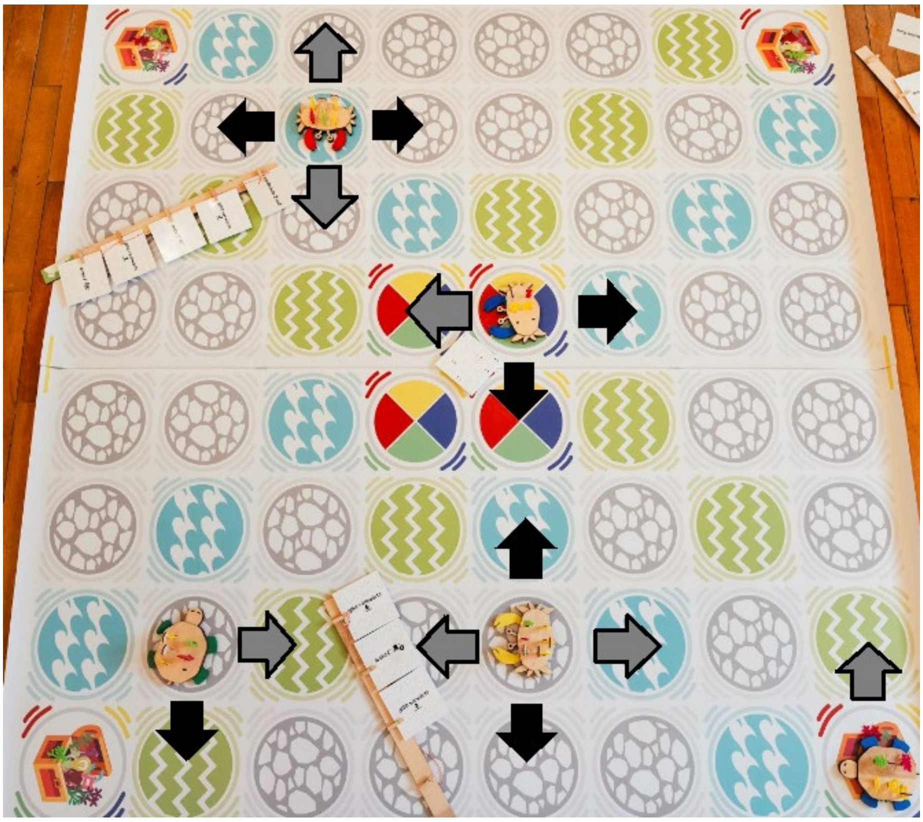

Figure 2. The Treasure Hunt game restrictions in movement for a certain scenario (black arrows: movements normally allowed for each game piece; grey arrows: bonus movements for each game piece due to special cards).

2.1.3 Scoring

While playing the game, teams collect inventory items that they carry on the back of their game pieces. These inventory items include food-treasure items and skill badges. There are 6 different food-treasure items, differentiated both by shape and colour. Green food items represent earth growing food, and magenta ones represent seaweed to-be-collected by turtles and crabs, respectively. There are 4 different skill badges (see Figure 3) that teams can collect at the end of each turn: (i) sequence-, (ii) loop-, (iii) angular degree-, and (iv) efficiency-badges. Teams get a sequence badge each time they succeed in forming and executing a correct algorithm reflecting a sequence of at least 3 commands. They get a loop or angular degree badge each time they succeed in using correctly and meaningfully a loop card or a turning card, respectively. Furthermore, in case teams build and execute correctly an algorithm consisting of a sequence of 5 commands, they get an efficiency badge because using 5 cards in an efficient combination is difficult. At the end of the game, scores are summed up based on both food-treasure items collected and on the number of skill badges. Winner of the game, however, is considered the team that first collects three food-treasure items.

2.2 The Race

The Race is the second game of *Crabs & Turtles.* In this game, users have to reach the end first by solving and handling math-related riddles and events. To do so, players have to manipulate constants and changing values of variables or make decisions based on conditionals. Overall, this game focuses more on CT abilities related to mathematical skills. In particular, understanding of constants and variables, operators, and events handling, addition, and multiplication, as well as simple conditional understanding, are the main abilities to be acquired or trained, respectively, in this game.

Figure 3. The Treasure Hunt inventory items (1. Food treasure for crabs, 2. Food treasure for turtles, 3. Sequence badge, 4. Angular degree badge, 5. Loop badge, 6. Sequence board, 7. Game piece, 8. Re-writable motion command cards).

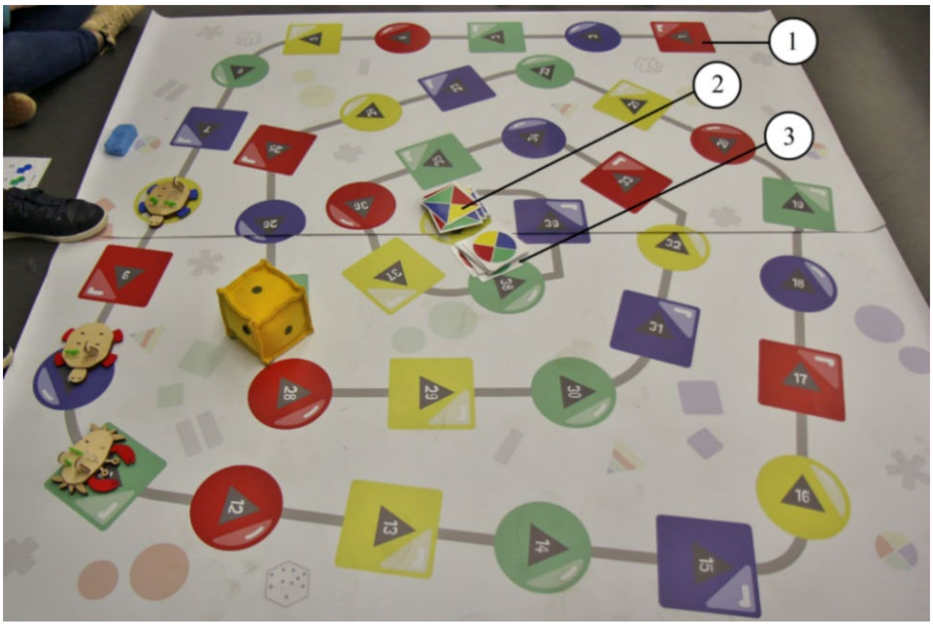

Figure 4. The Race game (1. Starting point, 2. Event cards, 3. Riddle cards).

2.2.1 Learning objectives

Main aim of the game is the general introduction of concepts shared between coding and math. Players are introduced to variables, constants, operators, and conditionals. The game focuses on training to handle values within simple and more complex arithmetic operations that contain additions, subtractions, and multiplications. Operations consist of visual representations of variables and constants (see Figure 5), so that the players get familiar with recognizing symbolic representations of things, get used to an abstract form of reading instructions, and handling events that allow for generalization, as it happens in actual coding.

The game is played in teams of two, and each team possesses one out of two types of game pieces, a re-writable variable/notes board, a marker, and a sponge. The game starts with all the game pieces placed at the starting point on the game board (see Figure 4). In each round, each team rolls a dice and moves as many steps forward as shown on the dice. In case players move their game piece to a circle shape, they receive a card from the pile of circle cards (Riddle cards); otherwise, they receive a card from the pile of square cards (Event cards). Circle cards contain riddles of equations that players have to solve. An example is presented in Figure 5 (lower panel), where the riddle asks the solution of an addition. The triangle represents the number of the step on which the game piece is currently standing, the colourful circle and square, respectively, indicate the value of the colour variable on which the game piece stands, and the third part of the addition is the value of the dice in the current turn. In the example depicted in Figure 5, assuming that the red crab of Figure 4 is playing and the team rolls the dice, and it shows 1, the crab will have to move forward one step. From the green square point number 11 it will move on one step to the red circle point number 12. The team will take a Riddle card and will have to solve the riddle. The addition consists of the value of the grey triangle, which is currently 12, the value of the red color variable that has been influenced by all previous turns so far and the value of the dice, which for this turn was 1.

2.2.2 Gameplay and rules

When the team solves the riddle correctly, they can move forward a defined number of steps, written on the lower right corner of the card (in the example of Figure 5 the steps forward are 3). The more complicated a riddle, the more steps the team is allowed to move forward. The difficulty of the riddles arises from the number of operators, variables, and constants in the mathematical expression (polynomial) of each riddle. For example, a mathematical expression that involves both variables and constants is considered more complicated than a mathematical expression consisting only from constants. Square cards, in contrast, contain events that change the value of a variable. These changes need to be calculated from simple numerical operations and/or conditionals specified on each square card. For example, in Figure 5, the Event card describes a conditional event, in which the player has to recalculate the value of the colour variable his/her game piece has landed on. When, for example, the game piece has landed on the red colour, and at this point of the game, the red colour variable has a value lower than 10, then the new value that the red colour variable will receive is calculated from a multiplication of the current value by 2. Otherwise, the new value of the variable would be the current one reduced by 9. When variable values are handled correctly, players move forward one step and wait for their next turn. To keep track of the changing colour variable values within the game, teams use the rewritable variable board (see Figure 6). It is also used as a re-writable note board for math calculations. The overall aim of this game is to reach the final point in the centre of the game board as fast as possible. While teams solve riddles and interact with events, they collect specific skill badges (see Scoring) that they collect on their game pieces. The first team to reach the centre of the spiral has to wait for the round to finish. In case there is another team to reach the end of the race within the same round, the winner is determined based on the number of badges collected.

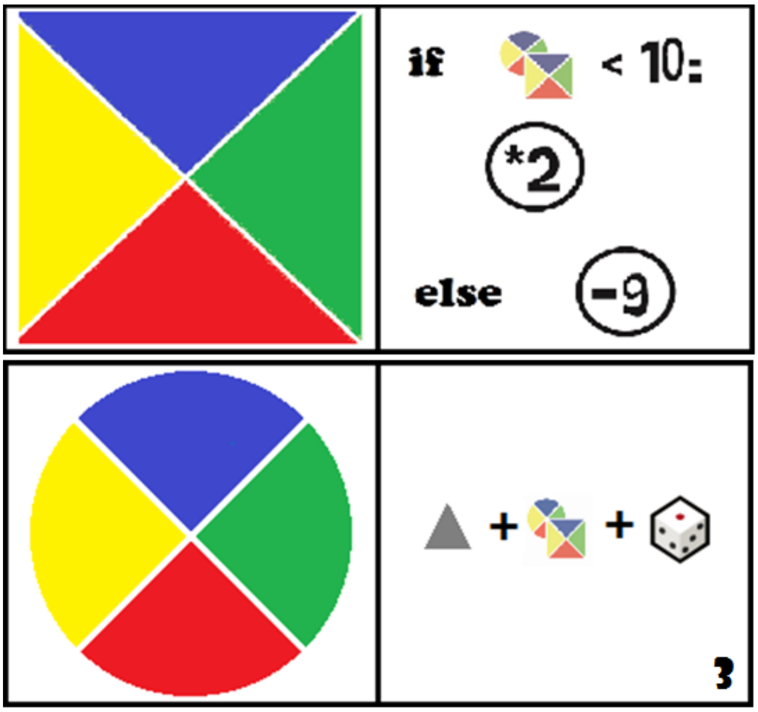

Figure 5. The Race game, an Event card example (upper panel, colourful square; left: cover, right: content) and a Riddle card example (lower panel, colourful circle; left: cover, right: content).

2.2.3 Scoring

Inventory items in *The Race* consist of 6 different skill badges: (i) addition-, (ii) subtraction-, (iii) multiplication-, (iv) variable-, (v) constant-, and (vi) conditional-badge. Teams get a variable or a constant badge each time they succeed in recognizing and handling a variable and a constant, respectively. They also get an addition, subtraction, or multiplication badge each time they solve the respective operation requested on a circle card correctly. Each time a team handles an Event card that indicated a value change of a conditional variable (see Figure 5, upper panel) correctly, it gets a conditional badge as well. At the end of the game, scores are summed up based on the inventory items of each team. The winner of the game is the first team to reach the end of the game board with the most badges collected.

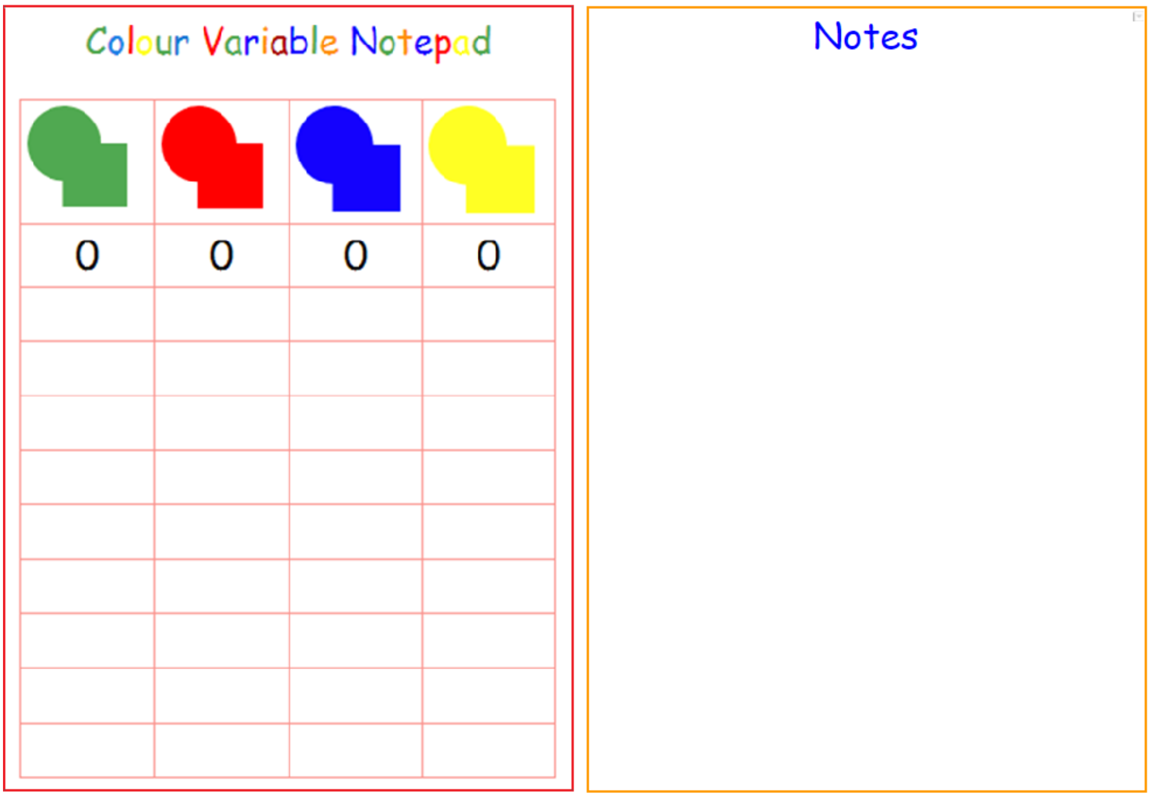

Figure 6. The Race game, the re-writable variable board.

2.3 Patterns

Patterns is a card game, played by individual players and not teams. In this game, players have to find patterns and match cards by certain rules, as fast as possible. This procedure is closely related to pattern recognition processes that are necessary for coding, for instance, when decomposing problems, generalizing solutions and forming loops.

2.3.1 Learning objectives

The game is an introductory activity to the concept of patterns. Patterns are crucial concepts in CT. They are used both in identifying abstractions and generalization (Curzon & McOwen, 2017). This game aims at training to recognize shape and colour patterns by following specific rules.

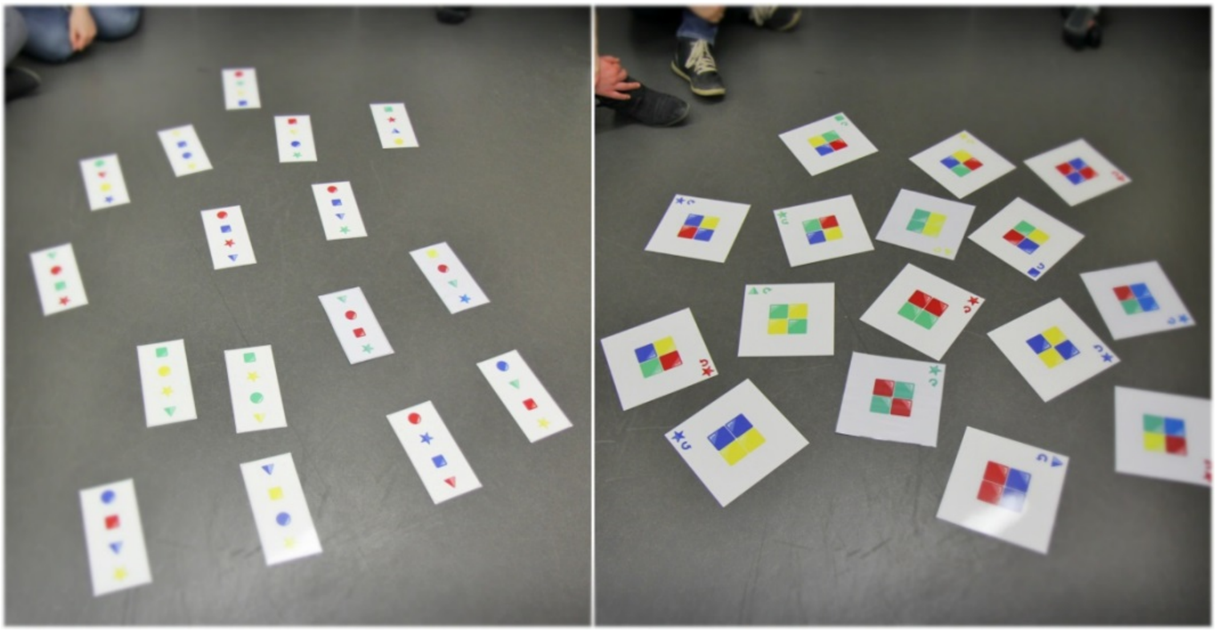

Figure 7. Patterns game.

2.3.2 Gameplay and rules

The game is structured in two parts and needs a game master that will act as the card dealer. Half of the cards lay open in arbitrary order on the floor, while the other half of the cards is being gradually revealed by the game master in random order. Players begin playing with no cards and start collecting cards each time they find a correct pair. In the first part (Figure 7, left), players have to find and match two cards according to the pattern depicted on them as fast as possible. In order for a player to claim a pair, he/she has to be the fastest in turn-taking from the game-master, by raising a hand. The game master reveals cards one after the other, and the first player to recognize and match two cards correctly wins the paired card and thus a point. When a card is revealed but not correctly paired, the game master hides it again and opens a new one. To match rectangle cards, players have to follow three rules: i. cards should share the same pattern order (e.g., star, triangle, circle, square, see Figure 8 left panel), ii. they should not have the same colour at the same position (e.g., in Figure 8, colour order of the matching card should not be blue-green-red-blue), and iii. cards should share the same colour palette (e.g., in Figure 8 only the colours blue, green, and red). In the second part of the game (Figure 8, right panel) players should read the shape and colour code of the square cards correctly and match them to the correct rectangle card (Figure 8 left) as fast as possible. The square card's code is read clockwise or counterclockwise, starting from the indicated shape and following the colour order directed by an arrow (i.e., starting with a blue star, going clockwise to green, red, and blue again, which matches the shape and colour order of the rectangle card, left panel). When players match a rectangle to a square card correctly, they win the square card.

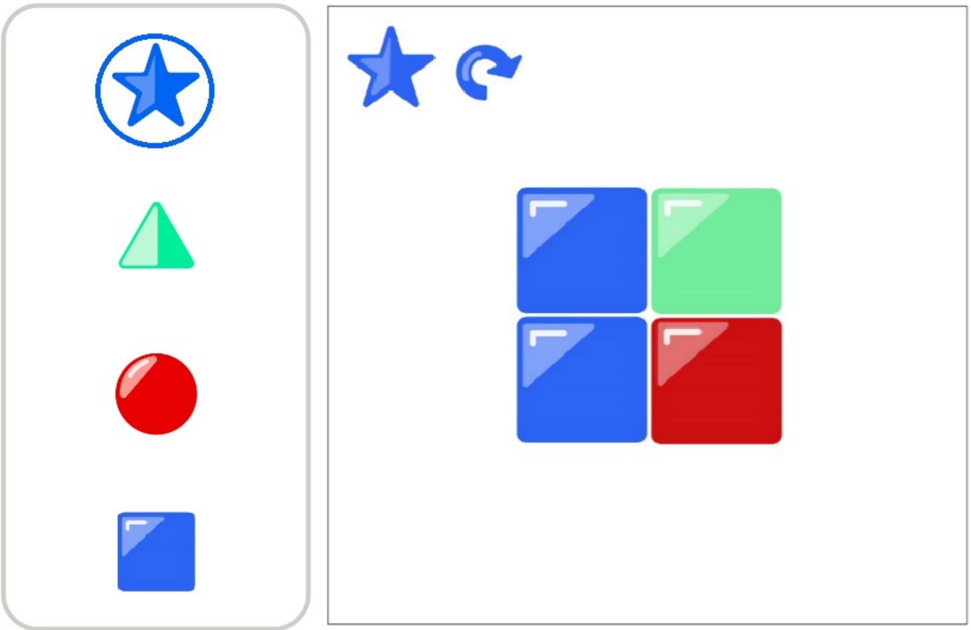

Figure 8. Patterns game cards; left panel: rectangle card; right panel: square card.

2.3.3 Scoring

In this game, players gain cards when they match cards in pairs correctly. When a pair is correctly paired, the player to claim it collects the respective cards from the floor. At the end of the game, players count the number of cards they collected while playing. The player who won the most cards wins the game.

3 Pilot Evaluation

Regular user tests are an important step during the development of a new game, educational or not. Here, we describe the results of a 2-phase user test evaluating the game experience of the above-described games. Our primary focus was on investigating game experience quantitatively and gathering qualitative feedback of participants to identify potential dysfunctionalities during game-play, which then can be addressed before testing it with our main target group of primary school children. The main objective of phase 1 was to evaluate the overall game experience of *Crabs and Turtles* as indicated by users. This was followed by a more in-depth analysis of game experience in phase 2, which investigated the games *The Treasure Hunt*, *The Race*, and *Patterns* subsumed in *Crabs and Turtles* separately. Moreover, to further validate the current approach, participants in phase 1 consisted of regular university students, while in phase 2, relevant stakeholders were tested: teachers, computer science instructors, and professional gamification experts.

3.1 Study phase 1

3.1.1 Participants

We collected data from 17 adult university students from the University of Tuebingen, aged between 22 and 33 (*mean* = 27.12, *SD* = 3.20). Students participated voluntarily.

3.1.2 Procedure and materials

In 3 separate gaming sessions (á ~2 hours), we evaluated the game experience of participants. In each session, all 3 games of *Crabs and Turtles* were played. After a short introduction to the aim of the session, participants were asked to fill in an optional photographic release form. Before participants started playing each of the games, we provided oral and visual instructions. After playing all 3 games, participants were asked to fill in the Game Experience Questionnaire-GEQ (Poels et al., 2007). We used the English version of the Core (33 items) and the Social Presence (17 items) modules to assess overall game experience. The Core module consists of seven subscales addressing i. Immersion, ii. Flow, iii. Competence, iv. Positive Affect, v. Negative Affect, vi. Tension and vii. Challenge. The Social Presence module consists of three subscales assessing i. Empathy, ii. Negative Feelings and iii. Behavioural Involvement. For each subscale, we used the average scores of the respective items as the dependent variable. Each item had to be responded to on a 5-point Likert-scale (1 = not at all; 2 = slightly; 3 = moderately; 4 = fairly; 5 = extremely). For example, the first item of the Core module reads as follows: *"I felt content",* and participants had to rate their experience of content on the aforementioned Likert scale by crossing an answer from 1 to 5 (e.g., crossing 2 would mean *"I felt slightly content"*).

Moreover, we used 4 additional items to evaluate further overall game experience, which also employed a 5-point Likert-scale: *Q1. I would explain my experience as playing*; *Q2*. *I would explain my experience as learning* (Q1 & Q2: 1 = not at all; 2 = not really; 3 = undecided; 4 = somewhat; 5 = very much); *Q3*. *I would recommend the games to a friend*; *Q4. I would like to play the games again in the future* (Q3 & Q4: 1 = not at all; 2 = not really; 3 = undecided; 4 = likely; 5 = very likely). Those 4 items were added to the questionnaire with the intention to measure the perception of the game as learning and/or playing, as the GEQ questionnaire aims at evaluating game experience more broadly and not game experience for educational games in particular.

Finally, 5 more items were used to evaluate specific design elements of the game, i.e., board, cards, game pieces, inventory items, and rules, again using a 5-point Likert-scale (1 = not at all; 2 = slightly; 3 = moderately; 4 = fairly; 5 = extremely). Moreover, every session included an open discussion part to gather qualitative feedback from the participants.

3.1.3 Results

Game experience: Mean values of GEQ subscales were considered to reflect game experience in this phase. We used a conservative approach of analyzing each subscale by conducting one-sample *t*-test comparing means of subscale ratings of the middle value of the scale (3 = mediocre) of the 5-point Likert scale. Descriptive results and inferential statistics of Core and Social Presence subscales are summarized in Table 3 below. Participants' ratings of the games on the *Competence*, *Sensory & Imaginative Immersion* and *Positive affect* subscales of the Core module were significantly higher than mediocre (see Table 3). In contrast, ratings on the *Tension/Annoyance*, *Challenge*, and *Negative Affect* subscales of the Core module and the *Negative Feelings* and *Behavioural Involvement* subscales of the Social Presence module were significantly lower than mediocre (see Table 3). We did not observe significant differences to mediocre for the subscale *Flow* of the Core module and the subscale *Empathy* of the Social Presence module.

Different from the GEQ data, for which we conducted *t*-tests on averaged ratings for the different subscales, differentiating aspects for overall experience and specific design elements meant conducting *t*-tests against mediocre on the data of individual items. Therefore, the respective results should be interpreted more cautiously.

Table 3. Mean scores for Core Module and Social Presence module of GEQ at phase 1.

GEQ modules	*mean*	*SD*	*df*	*t*	*p*	*α*	*α**
Core							
Competence	3.72	.90	16	3.30	0.005	.913	.826
Sensory & Imaginative Immersion	3.63	.73	16	3.59	0.002	.783	.891
Flow	3.02	.83	16	0.12	0.908	.844	.866
Tension/Annoyance	1.29	.37	16	-18.99	0.001	.402	.811
Challenge	2.38	.65	16	-3.95	0.001	.735	.745
Negative affect	1.84	.53	16	-9.04	0.001	.456	.712
Positive affect	4.19	.46	16	10.76	0.001	.770	.797
Social Presence							
Empathy	3.27	.60	16	1.89	0.077	.779	.886
Negative Feelings	2.08	.73	16	-5.17	0.001	.876	.860
Behavioural Involvement	2.47	.73	16	-3.00	0.008	.829	.711

Overall experience: Participants perceived their experience as somewhat playing (Q1: *mean* = 4.18, *SD* = .73; $t(16)$ = 6.67, $p < 0.001$) as indicated by a rating significantly above "undecided", but not so much as a learning experience (Q2: *mean* = 3.35, *SD* = 1.22; $t(16)$ = 1.19, p = 0.251). Participants reported that they would likely to very likely recommend the game to a friend (Q3: *mean* = 4.53, *SD* = .72; $t(16)$ = 8.79, $p < 0.001$) and would likely play the game again in the future (Q4: *mean* = 4.18, *SD* = .64; $t(16)$ = 7.63, $p < 0.001$) as reflected by ratings significantly above "undecided".

Evaluation of specific design elements: The five different design elements measured by the questionnaire scored a mean of 4.46 (*SD* = .44) on the 5-point Likert scale. More specifically, users rated each design element (Board: *mean* =4.5, *SD* =.61, $t(16)$ = 10.10, $p < 0.001$; Cards: *mean* = 4.38, *SD* = .70, $t(16)$ = 8.15, $p < 0.001$; Game pieces: *mean* = 4.88, *SD* = .33, $t(16)$ = 23.38, $p < 0.001$; Inventory items: *mean* = 4.19, *SD* = .81, $t(16)$ = 6.06, $p < 0.001$, and Rules: *mean* = 4.38, *SD* = .78, $t(16)$ = 7.26, $p < 0.001$) significantly above mediocre. The ratings of one participant were missing for the design module of the questionnaire and were replaced with the mean of the sample.

Internal consistency of GEQ, as reported by (Poels et al., 2007) and reflected by Cronbach's alpha, is given in Table 3 (column *α**). Besides, Cronbach's alpha, as obtained in the current study, is also reported in Table 3 (column *α*). The observed internal consistency indicated acceptable reliability for most subscales with $\alpha > .70$. However, this was not the case for subscales *Tension/Annoyance* and *Negative affect* of the GEQ Core module. *Qualitative feedback:* Participants' impressions on the games were positive and encouraging. Their comments in phase 1 led to several design changes regarding the game mechanisms. For instance, it became clear that instructions were not always evident in how they were presented to players. Moreover, participants reported some in-game unbalances caused by a high dependency on chance. For example, in *The Treasure Hunt,* the command card for loops was part of the Event cards which are taken at each turn by chance from the pile. That was affecting teams' strategy to prepare their moves. Therefore, in phase 2, this card was given as a Motion command card to each team from the beginning of the game. Reported unbalances of this kind were adjusted by excluding and/or adding specific kinds of cards (i.e., in *The Treasure Hunt,* in *The Race*) in an effort to balance chance and skill-driven strategies during gameplay. Participants also criticized other dysfunctionalities, like long waiting in between turns or very limited step movement. We addressed this by introducing a time limit for each turn using a 3-minute hourglass and by excluding the use of a dice for determining the number of steps allowed for a team to move per turn (i.e., in *The Treasure Hunt*). These problems were addressed and fixed before starting the assessment of each game separately in phase 2.

3.2 Study phase 2

3.2.1 Objective and Procedure

The second phase of the adult sessions aimed at a more in-depth evaluation of game experience by investing each game separately. This phase also consisted of 3 sessions (á 2-3 hours), one hosted by the 11th Thessaloniki Gamification Meet-up, and two independently organized events. The procedure followed in this phase was comparable to that of phase 1 with the difference that participants had to fill in the respective questionnaires separately for each game.

3.2.2 Participants

Data were collected from 19 participants in total, aged between 25 and 52 (*mean* = 31.43, *SD* = 6.17). There were 10 female and 9 male participants, including teachers, computer science instructors, professional gamification designers, etc. Due to technical and organizational problems, not all participants were able to play all three games. That is, 15 played all three games, 2 participants played only the first game, 1 participant played only games 1 and 2, and 1 participant only games 2 and 3.

3.2.3 Materials

We used the same questionnaires and additional items as in phase 1. However, participants had to answer the questions for each game separately. The questionnaire was completed in its original English version from 13 participants. Six participants felt more confident completing it in its Greek translation.

3.2.4 Results

We applied the same conservative approach of analyzing GEQ Core and Social Presence subscales. However, in phase 2, we were able to conduct the analyses for each game. Descriptive results and inferential statistics are reported in Table 4. There were few missing values, which were replaced by the mean score for the respective item computed from the other participants.

Game experience - The Treasure Hunt: Participants rated this game significantly above mediocre on the subscales *Competence*, *Sensory & Imaginative Immersion*, and *Positive Affect* of the Core module and *Empathy* of the Social Presence module. In contrast, ratings were significantly below mediocre for the subscales *Tension/Annoyance*, *Challenge*, and *Negative Affect* of the Core module and the *Negative Feelings* and *Behavioural Involvement* subscales of the Social Presence module. We did not find a significant difference from mediocre for the Flow subscale of the Core module.

Game experience - The Race: Participants' ratings for this game were significantly above mediocre for the *Competence*, *Sensory & Imaginative Immersion*, and *Positive affect* subscales of the Core module and the *Empathy* subscale of the Social Presence module. Contrarily, participants rated the game significantly below mediocre on the subscales *Tension/Annoyance*, *Challenge*, and *Negative Affect* of the Core module and the *Negative Feelings* subscale of the Social Presence module. Again, we did not find a significant difference in the Flow subscale of the Core module and the *Behavioural Involvement* subscale of the Social Presence module.

Table 4. Mean scores for Core module and Social Presence module of GEQ at phase 2, per game-based activity.

			The Treasure Hunt				
GEQ modules	***mean***	***SD***	***df***	***t***	***p***	***α***	***α****
Core							
Competence	3.70	.75	17	3.98	0.001	.794	.826
Sensory & Imaginative Immersion	4.04	.71	17	6.19	0.001	.816	.891
Flow	3.17	.92	17	0.77	0.454	.764	.866
Tension/Annoyance	1.51	.48	17	-13.19	0.001	.130	.811
Challenge	2.21	.54	17	-6.17	0.001	.576	.745
Negative affect	1.42	.37	17	-17.97	0.001	.300	.712
Positive affect	4.29	.51	17	10.63	0.001	.800	.797
Social Presence							
Empathy	3.48	.88	17	2.32	0.033	.885	.886
Negative Feelings	2.04	.87	17	-4.69	0.001	.795	.860
Behavioural Involvement	2.49	.92	17	-2.36	0.030	.835	.711
			The Race				
GEQ modules	***mean***	***SD***	***df***	***t***	***p***	***α***	***α****
Core							
Competence	3.64	.88	16	2.96	0.009	.635	.826
Sensory & Imaginative Immersion	3.67	.86	16	3.21	0.005	.858	.891
Flow	3.28	1.14	16	1.02	0.324	.924	.866
Tension/Annoyance	1.84	1.09	16	-4.39	0.001	.830	.811
Challenge	2.42	.73	16	-3.24	0.005	.657	.745
Negative affect	1.69	.67	16	-8.05	0.001	.639	.712
Positive affect	4.14	.78	16	6.01	0.001	.894	.797
Social Presence							
Empathy	3.69	.84	16	3.38	0.004	.862	.886
Negative Feelings	2.26	.69	16	-4.44	0.001	.589	.860
Behavioural Involvement	2.55	1.06	16	-1.76	0.097	.876	.711
			Patterns				
GEQ modules	***mean***	***SD***	***df***	***t***	***p***	***α***	***α****
Core							
Competence	3.70	.90	15	3.09	0.007	.820	.826
Sensory & Imaginative Immersion	3.84	.72	15	4.66	0.001	.849	.891
Flow	3.90	.82	15	4.41	0.001	.772	.866
Tension/Annoyance	1.73	.84	15	-6.08	0.001	.675	.811
Challenge	3.18	.80	15	0.87	0.397	.706	.745
Negative affect	1.32	.51	15	-13.21	0.001	.487	.712
Positive affect	4.41	.65	15	8.68	0.001	.785	.797
Social Presence							
Empathy	3.27	.93	15	1.47	0.162	.840	.886
Negative Feelings	2.08	1.03	15	-3.47	0.003	.821	.860
Behavioural Involvement	2.47	1.28	15	-1.11	0.285	.933	.711

Games experience – Patterns: Largely similar to the results for the other games, participants' ratings for *Patterns* were significantly above mediocre for the *Competence*, *Sensory & Imaginative Immersion*, and *Positive affect* subscales of the Core module and the *Empathy* subscale of the Social Presence module. Again, ratings for the *Tension/Annoyance*, *Challenge*, and *Negative Affect* subscales of the Core module and the *Negative Feelings* and *Behavioural Involvement* subscales of the Social Presence module were significantly below mediocre. Also, we did not find a significant difference to mediocre for the subscale Flow of the Core module as well as the subscales *Empathy* and *Behavioural Involvement* of the Social Presence module.

Again, analysis of overall experience and specific design elements required us to run *t*-tests on data from individual items. Thus, the respective results should be interpreted more cautiously.

Overall Experience: Participants experienced *The Treasure Hunt* very likely as playing (Q1: *mean* = 4.72, *SD* = .46; $t(17) = 15.85$, $p < 0.001$) and somewhat as learning (Q2: *mean* = 4.22, *SD* = .94; $t(17) = 5.50$, $p < 0.001$) as reflected by ratings significantly above "undecided". Additionally, participants reported that they would very likely recommend the game to a friend (Q3: *mean* = 4.50, *SD* = .71; $t(17) = 9.00$, $p < 0.001$), and would likely play the game again in the future (Q4: *mean* = 4.22, *SD* = 1.06; $t(17) = 4.89$, $p < 0.001$), which was also supported by ratings significantly above "undecided".

For *The Race* ratings significantly above "undecided" indicated that participants rated their game experience very likely as playing (Q1: *mean* = 4.53, *SD* = .72; $t(16) = 8.79$, $p < 0.001$), and somewhat as learning (Q2: *mean* = 4.24, *SD* = 1.03; $t(16) = 4.93$, $p < 0.001$). Furthermore, ratings significantly above "undecided" substantiated that they would likely recommend the game to a friend (Q3: *mean* = 4.18, *SD* = .73; $t(16) = 6.67$, $p < 0.001$), and also would likely play it again in the future (Q4: *mean* = 4.29, *SD* = .69; $t(16) = 7.78$, $p < 0.001$).

Finally, participants perceived the *Patterns* game very likely as a playing experience (Q1: *mean* = 4.69, *SD* = .79; $t(15) = 8.51$, $p < 0.001$) and likely as learning (Q2: *mean* = 4.06, *SD* = 1.00; $t(15) = 4.26$, $p < 0.001$), which was again reflected by ratings above "undecided". Moreover, according to ratings above "undecided", participants reported that they would very likely recommend it to a friend (Q3: *mean* = 4.56, *SD* = .62; $t(15) = 9.93$, $p < 0.001$), and also play it again (Q4: *mean* = 4.56, *SD* = .73; $t(15) = 8.60$, $p < 0.001$).

Evaluation of specific design elements: In the current phase, the five design elements were evaluated individually. For the two first games all five design elements were assessed. The design elements of Games 1 and 2 scored a mean of 4.37 (*SD* = .47) and 4.03 (*SD* =.72) respectively, on the 5-point Likert scale. Game 3 design elements were scored a mean of 4.49 (*SD* =.51). More specifically, users liked all five design elements of *The Treasure Hunt* (Board: *mean* =4.25*, SD* =.94, $t(17)$ = 5.65, p < 0.001; Cards: *mean* = 4.13*, SD* = .83, $t(17)$ = 5.74, p < 0.001; Game Pieces: *mean* = 4.80, *SD* = .38, $t(17)$ = 20.33, p < 0.001; Inventory items: *mean* = 4.63, *SD* = .58, $t(17)$ = 11.86, p < 0.001 and Rules: *mean* = 4.06, *SD* = 1.00, $t(17)$ = 4.52, p < 0.001) significantly above than mediocre. Same positive scores received all the five design elements of *The Race* (Board: *mean* =3.69*, SD =1.3*, $t(16)$ = 2.17, p = 0.046; Cards: *mean* = 4.19*, SD* = 1.01, $t(16)$ = 4.83, p < 0.001; Game pieces: *mean* = 4.38, *SD* = .93, $t(16)$ = 6.12, p < 0.001; Inventory items: *mean* = 4.13, *SD* = .93, $t(16)$ = 5.01, p < 0.001 and Rules: *mean* = 3.75, *SD* = 1.15, $t(16)$ = 2.70, p = 0.016) as reflected by ratings significantly above mediocre. The two design elements in the questionnaire for *Patterns* scored also positively) as indicated by ratings significantly above mediocre (Cards: *mean* = 4.38*, SD* = .72, $t(15)$ = 7.65, p < 0.001; and Rules: *mean* = 4.60, *SD* = .49, $t(15)$ = 13.06, p < 0.001). The missing values for this part of the questionnaires were managed as before. There were two participants that did not fill in all the five questions concerning the design evaluation of the first game. Their missing values, as well as one single missing value from a third participant, were replaced by the mean scores of each single item. In the second game, one participant's responses were missing for all the 5 design elements.

For phase 2, reliability analyses run separately for each GEQ subscale game again indicated acceptable reliability for most subscales with $\alpha > .70$ (see Table 4, column α). However, this was not case for the subscales *Competence* (in *The Race*), *Tension/Annoyance* (in *The Treasure Hunt* and *Patterns*), *Challenge* (in *The Treasure Hunt* and *The Race*), *Negative affect* (*The Treasure Hunt*, *The Race,* and *Patterns*) and *Negative feelings* (in *The Race*). Those results may be affected by the rather small number of participants in our sample.*Qualitative feedback:* Participants' impressions of the games were positive and promising for the content and the mode of the games. Their comments in phase 2 were taken into consideration and led to minor changes in the latest version of the games. For instance, in *The Treasure Hunt,* the maximum duration of play during a turn (3 minutes) was considered too long; thus it was reduced and limited to 1 minute. In *The Race,* the depiction of variables on the cards was somewhat confusing; for that reason, the image was slightly adjusted. In *Patterns,* several participants (2 of them with partial colour blindness) reported colour confusion while trying to recognize shapes of yellow colour. Consequently, we changed the hue of yellow colour on the cards of this game. Finally, many participants requested a cumulative score across all three games that would allow determining an overall winner of *Crabs & Turtles*.

4 Discussion

The main aim of the present study was to describe the design and development of three unplugged games to foster computational thinking abilities in primary school children. The three games focused on different concepts relevant to computational thinking. In a 2-phase process, we evaluated users' game experience. Using an iterative user-centred development process, dysfunctionalities in gameplay and shortcomings in instructions were identified and fixed during the development process. Quantitative analyses of overall (phase 1) and game-specific game experience (phase 2) provided promising results as to the validity of our approach. In the following, we will discuss the results of phases 1 and 2 in turn.

In phase 1, university student participants rated their overall game experience after playing all three games. Results indicated an overall positive reception of the educational games. In particular, users reported to feel competent and immersed during gameplay and perceived positive affect. In contrast, their GEQ ratings did neither indicate the experience of tension nor did they report to perceive negative emotions more generally. Additional analyses of overall experience further indicated that the games were primarily perceived as a playful activity and only to a lesser degree as learning. These results are in line with our objective of conveying basic concepts of computational thinking in a low threshold and game-based manner. Importantly, this is also reflected in users reported willingness to play the games again and also recommend playing the respective games to friends. Therefore, the overall evaluation of the games yielded promising results about users' game experience that further backed the design of *Crabs &Turtles* as a whole.

In the more in-depth analysis of each individual game in phase 2 results of phase 1 were substantiated as we identified similar patterns for participants' ratings of game experience. Importantly, participants with a more educational oriented background (i.e., teacher, computer science instructors, etc.) again indicated that they perceived high levels of positive emotions, competence, as well as immersion while playing each of the three games.

Additionally, they reported only low perceived levels of negative emotions and tension in all three games. Interestingly, high levels of flow were only reported by participants for playing *Patterns*. At the same time, the overall challenge was rated relatively low in *The Treasure Hunt* and *The Race*, suggesting that these games in their current form might be rather easy for adult participants. This might also explain the rather mediocre perception of flow in these two games. Nevertheless, all three games were perceived as playful activities, and users indicated that they would like to play again as well as recommend all of the games to their friends. Moreover, the design elements of each game (i.e., game board, cards, game pieces, inventory items, rules) were rated positively throughout.

Taken together, results of phase 1, as well as phase 2 evaluations, provided converging evidence on the validity of *Crabs &Turtles* as an unplugged and game-based approach to convey basic concepts of computational thinking – both overall (phase 1) but also when considered separately for the three games *The Treasure Hunt*, *The Race*, and *Patterns*.

As such, this indicated that we took the first steps in developing an educational game. The design and CT concepts employed in all three games were derived from recent research (Berland & Lee, 2011; Brennan & Resnick, 2012a; Weintrop, Holbert, et al., 2016), which is a first crucial step in developing educational (board) games. From the beginning, we used an iterative user-centred development procedure, starting with pilot tests with primary school children, our main target group. However, before starting a comprehensive evaluation of cognitive effects and learning outcomes due to the three games in our main target group, we aimed at optimizing game experience and in-game procedures. Therefore, we employed a 2-phase evaluation of game experience in adults, as reported in the current article. Overall and specific game experience was consistently positive, as indicated by participants' ratings in both phases. Moreover, qualitative feedback by users helped to further optimize gameplay and mechanics. Adult participants, in particular the specialized group of teachers, computer science instructors, and gamification experts considered in phase 2, was able to provide us with specific and to-the-point feedback to further develop and improve the games and prepare them for the use in our main target group.

Based on our observations, we are confident that the employed alternation of different mechanics across the games helped to keep different personal characters (e.g. shy, extrovert, patient or impatient participant) engaged in the educational content of the game. We noted that for the mechanic of turns, seemingly more analytical and patient individuals with a focus on details seemed to be attracted more to the gameplay in the first game (*The Treasure Hunt*). In contrast, in the third game (*Patterns*), we found that seemingly more impatient and extrovert users were highly engaged. For the second game (*The Race*), the mechanism of turns seemed to engage all kinds of users more equally after modification of the rules from phase 1 to phase 2. Another supporting example is the alternating mode of gameplay across games, starting with cooperation within teams and competition between teams (*The Treasure Hunt* and *The Race*) moving on to the final game with competition between all participating individuals (*Patterns*). This alternation of modes across games supported the collaborative introduction to the games (*The Treasure Hunt* and *The Race*), as well as the personal satisfaction of each user at the endmost of the games (*Patterns*).

There are limitations to the present study that need to be considered when interpreting the results and should be addressed in already planned follow-up studies to overcome these limitations. For instance, while adult participants might be well able to provide more specific feedback and are easier to recruit and test as compared to children for initial pilot tests, a comprehensive analysis of game experience and learning outcomes in the main target group is, of course, necessary. Therefore, such a comprehensive evaluation of our games will be our next step, with a special focus on learning outcomes (by implementing a pre-/post-test design) in addition to questionnaire data on game and learning experience.

5 Perspectives

Future studies will, thus, have to evaluate game experience but also learning outcomes of the three games in primary school children to appraise their educational value in fostering CT abilities. Moreover, these games will be integrated into the first three lessons of a 10 lesson CT course curriculum (Tsarava et al., 2017). In this CT course, a game-based introduction of CT concepts in an unplugged manner is provided (i.e., without using a computer or other digital technology). In later lessons of the course, the very same CT concepts are picked up again in the context of other educational programming environments, for instance, Scratch, Scratch for Arduino (S4A), and Roberta robot programming (for a more comprehensive description of the course curriculum see Tsarava et al., 2017). Generally, the extra-curricular course primarily aims at introducing and fostering computational thinking, but not exclusively in gifted students between 7 and 9 years of age. More specifically, in a first phase, the games described in this article are planned to be evaluated in 4 Academies of the Hector Children's Academy Program (HCAP) for gifted children, as one of 10 Hector Core Courses developed by the Hector Research Institute of Education Sciences and Psychology in Germany. In a second phase, the course curriculum, including the three life-size board games, will be taught in more than twenty Hector Children's Academies across Baden-Wuerttemberg, Germany.

Besides an overall evaluation of the educational value of the 3 games presented in terms of learning outcomes, we will specifically investigate whether game metrics, such as acquired badges and points, may provide a valid and reliable stealth assessment tool to allow for formative assessment of CT abilities (Shute & Kim, 2014). Finally, this upcoming comprehensive evaluation aims at investigating the underlying cognitive abilities involved in CT and possible transfer effects of the training by administering standardized psychological tests to allow for a differential view on CT (Shute et al., 2017).

References

Apostolellis, P., Stewart, M., Frisina, C., & Kafura, D. (2014). RaBit EscAPE: A Board Game for Computational Thinking. *Proceedings of the 2014 Conference on Interaction Design and Children - IDC '14*, 349–352. https://doi.org/10.1145/2593968.2610489

Astrachan, O., & Briggs, A. (2012). *The CS Principles Project*. *3*(2), 38–42.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(August), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bauer, A., Butler, E., & Popovic, Z. (2015). Approaches for teaching computational thinking strategies in an educational game: A position paper. *Proceedings - 2015 IEEE Blocks and Beyond Workshop, Blocks and Beyond 2015*, 121–123. https://doi.org/10.1109/BLOCKS.2015.7369019

Berland, M., & Lee, V. R. (2011). Collaborative Strategic Board Games as a Site for Distributed Computational Thinking. *International Journal of Game-Based Learning*, *1*(2), 65–81. https://doi.org/10.4018/ijgbl.2011040105

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers and Education*, *94*, 178–192. https://doi.org/10.1016/j.compedu.2015.11.003

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. *Annual American Educational Research Association Meeting, Vancouver, BC, Canada*, 1–25. http://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf

Brown, N. C. C., Sentance, S. U. E., Crick, T. O. M., & Humphreys, S. (2014). Restart: The Resurgence of Computer Science in UK Schools. *ACM Transactions on Computing Education*, *14*(2), 1–22. https://doi.org/10.1145/2602484

Butz, M. V. (2016). Toward a unified sub-symbolic computational theory of cognition. *Frontiers in Psychology*, *7*(JUN), 1–19. https://doi.org/10.3389/fpsyg.2016.00925

Code.org. (n.d.). https://code.org/

Curzon, P., & McOwen, P. W. (2017). *The power of Computational Thinking* (1st ed.). World Scientific.

Dan, S. (2013). *Robot Turtles*. http://www.robotturtles.com/

Echeverría, A., García-Campo, C., Nussbaum, M., Gil, F., Villalta, M., Améstica, M., & Echeverría, S. (2011). A framework for the design and integration of collaborative classroom games. *Computers and Education*, *57*(1), 1127–1136. https://doi.org/10.1016/j.compedu.2010.12.010

European Coding Initiative. (n.d.). http://www.allyouneediscode.eu/

European School Network. (n.d.). https://www.esnetwork.eu/

Fraser, N. (2015). Ten things we've learned from Blockly. *Proceedings - 2015 IEEE Blocks and Beyond Workshop, Blocks and Beyond 2015*, 49–50. https://doi.org/10.1109/BLOCKS.2015.7369000

Fullerton, T. (2008). Game Design Workshop: A Playcentric Approach to Creating Innovative Games. In *Technology*. Elsevier Inc. https://doi.org/10.1007/s13398-014-0173-7.2

Garcia-Peñalvo, F. J. (2016). What Computational Thinking Is. *Journal of Information Technology*

Research, *9(3)*, *v–vi*(October).

Grover, S., & Pea, R. (2013). Computational Thinking in K-12: A Review of the State of the Field. *Educational Researcher*, *42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

Kafai, Y. B., & Burke, Q. (2015). *Constructionist Gaming : Understanding the Benefits of Making Games for Learning*. *50*(4), 313–334. https://doi.org/10.1080/00461520.2015.1124022

Kazimoglu, C. (2013). *Empirical evidence that proves a serious game is an educationally effective tool for learning computer programming constructs at the computational thinking level*. *June*.

Kazimoglu, C., Kiernan, M., Bacon, L., & Mackinnon, L. (2012). A Serious Game for Developing Computational Thinking and Learning Introductory Computer Programming. *Procedia - Social and Behavioral Sciences*, *47*, 1991–1999. https://doi.org/10.1016/j.sbspro.2012.06.938

Leacock, M. (2012). *Pandemic*. http://www.leacock.com/games/#/pandemic/

McKinley, S. R. (2006). *Qwirkle*. Mindware.

National Science Foundation. (n.d.). https://www.nsf.gov/

Osmo Coding Family. (2016). https://www.playosmo.com/en/coding-family/

Papert, S. (1994). Philosophy and the Computer. *Philosophical Books*, *35*(1), 39–41. https://doi.org/10.1111/j.1468-0149.1994.tb02396.x

Papert, S., & Solomon, C. (1971). *Twenty Things To Do With a Computer*.

Pea, R. D., & Kurland, D. M. (1984). On the cognitive effects of learning computer programming. *New Ideas in Psychology*, *2*(2), 137–168. https://doi.org/10.1016/0732-118X(84)90018-7

Plass, J. L., Homer, B. D., Kinzer, C. K., Plass, J. L., Homer, B. D., Kinzer, C. K., Plass, J. L., Homer, B. D., & Kinzer, C. K. (2016). *Foundations of Game-Based Learning Foundations of Game-Based Learning*. *1520*(February). https://doi.org/10.1080/00461520.2015.1122533

Poels, K., de Kort, Y., & Ijsselsteijn, W. (2007). *FUGA - The fun of gaming: Measuring the human experience of media enjoyment. Deliverable D3.3: Game Experience Questionnaire*.

Randolph, A. (1999). *Ricochet Robots*.

Roungas, B., & Dalpiaz, F. (2016). A Model-Driven Framework for Educational Game Design. *Revised Selected Papers of the 4th International Conference on Games and Learning Alliance - Volume 9599*, 1–11. https://doi.org/10.1007/978-3-319-40216-1_1

Shute, V. J., & Kim, Y. J. (2014). Formative and Stealth Assessment. *Handbook of Research on Educational Communications and Technology: Fourth Edition*, 311–321. https://doi.org/10.1007/978-1-4614-3185-5

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

Tsarava, K. (2016). *Programming in Greek with Python* [Aristotle Univerity of Thessaloniki]. http://ikee.lib.auth.gr/record/284234/files/GRI-2016-17183.PDF

Tsarava, K., Moeller, K., Pinkwart, N., Butz, M. V., Trautwein, U., & Ninaus, M. (2017). Training computational thinking: Game-based unplugged and plugged-in activities in primary school. *Proceedings of the 11th European Conference on Games Based Learning, ECGBL 2017*, *October*, 687–695.

Tuomi, P., Multisilta, J., Saarikoski, P., & Suominen, J. (2018). Coding skills as a success factor for a society. *Education and Information Technologies*, *23*(1), 419–434. https://doi.org/10.1007/s10639-017-9611-4

Wang, P. S. (2015). *From Computing to Computational Thinking* (1st ed.). Chapman and Hall/CRC.

Weintrop, D., Holbert, N., Horn, M. S., & Wilensky, U. (2016). Computational Thinking in Constructionist Video Games. *International Journal of Game-Based Learning*, *6*(1), 1–17. https://doi.org/10.4018/IJGBL.2016010101

Wing, J. M. (2006). Computational Thinking. *Theoretical Computer Science*, *49*(3), 33–35. https://doi.org/https://www.cs.cmu.edu/~15110-s13/Wing06-ct.pdf

Wing, J. M. (2010). Computational Thinking: What and Why? *Thelink - The Magaizne of the Varnegie Mellon University School of Computer Science*, *March 2006*, 1–6. http://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. https://doi.org/10.1037/a0031311

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational Thinking in Elementary and Secondary Teacher Education. *ACM Transactions on Computing Education*, *14*(1), 1–16. https://doi.org/10.1145/2576872

Board Games for Training Computational Thinking

Katerina Tsarava, Korbinian Moeller, Manuel Ninaus

Abstract

Computational thinking (CT) is a term widely used to describe algorithmic thinking and logic reasoning concepts and processes often related to computer programming. As such, CT as a cognitive ability, builds on concepts and processes that derive from computer programming but are applicable to wider real-life problems and STEM domains. CT has recently been argued to be a fundamental skill for 21st-century education and an early academic success indicator that should be introduced and trained already in primary school education. Accordingly, we developed three life-size board games – *Crabs & Turtles: A Series of Computational Adventures* – that aim at providing an unplugged, gamified and low-threshold introduction to CT by presenting basic coding concepts and computational thinking processes to 8 to 9-year-old primary school children. For the design and development of these educational board games, we followed a rapid prototyping approach. In the current study, we report the results of an empirical evaluation of the game experience of our educational board games with students of the target age group. In particular, we conducted quantitative analyses of player experience of primary school student participants. Results indicate an overall positive game experience for all three board games. Future studies are planned to further evaluate learning outcomes in educational interventions with children.

Keywords: computational thinking, unplugged activities, board games

1 Introduction

Computational Thinking (CT) denotes the mental ability of creating a computational solution to a problem, by first decomposing it, and then developing a structured and algorithmic solution procedure (Wing, 2006a, 2010). CT, as a cognitive ability, is argued to reflect the application of fundamental concepts and reasoning processes that derive from computer science and informatics to wider everyday life activities and problems but also STEM (Science, Technology, Engineering, and Mathematics) domains (Wang, 2015). The construct of CT as a cognitive ability shares common concepts with computer programming as a practical skill. Central concepts in computer programming are the ideas of sequences, operators, data/variables, conditional, events, loops, and parallelism (Brennan & Resnick, 2012a). Respectively, CT draws on processes such as decomposition, algorithmic thinking, conditional logic, pattern recognition, evaluation, abstraction, and generalization, which reflect cognitive counterparts of central computer programming concepts (Astrachan & Briggs, 2012; Wing, 2010).

CT, as a rather general problem-solving strategy applied to different domains, has been identified as a fundamental 21^{st}-century skill (Wing, 2006a). It has been suggested that the instruction on CT concepts may improve students' analytical skills and provide early indication and prediction of academic success (Haddad & Kalaani, 2015). Therefore, CT is considered a key competence for everyone and not just computer scientists (Wing, 2006a), comparable to literacy and numeracy (Yadav et al., 2014), that should be taught and acquired early in education.

Recent research focused on the benefits of CT and its integration into educational curricula, which has lately led to several adaptations and reformations of educational programs throughout all levels of education worldwide (Brown et al., 2014; Tuomi et al., 2018). Educational initiatives and governmental institutions all over the world have been working on the integration of CT into curricula of educational programs of primary, secondary, and higher education (*Code.Org*; *European Coding Initiative*; *European School Network*; *National Science Foundation*).

The societal relevance of CT led us to design and develop a CT training course for primary school children, introducing computer programming concepts and CT processes, applied to different STEAM (Science, Technology, Engineering, Art, and Mathematics) domains (for information on the overall course structure see Tsarava et al., 2017). Importantly, to offer a low threshold introduction to CT utilizing embodied learning (Barsalou, 2008), we developed unplugged life-size board games *Crabs & Turtles: A Series of Computational Adventures* (for a more detailed description of the games see Tsarava et al., 2018) for our CT training course.

Crabs & Turtles share common ideas with concepts of Papert's educational Logo Turtle (Papert, 1999) and logo-inspired gamified educational activities (Papert & Solomon, 1971). Logo Turtle transferred to the real world conceptualized ideas of programing-like commands and algorithms by applying them for the first time to a transparent moving and haptic object, the Turtle. The unplugged life-size game design allows embodied training (for the concept of embodied cognition, see Barsalou, 2008) of simple computational concepts and encourages active engagement and participation of students (for an overview, see Echeverría et al., 2011). The games' target group are primary and secondary school students (8-12 years old) with no prior programming knowledge. We deliberately chose an unplugged mode of the game, taking into consideration common concerns regarding the introduction of computer programming to young children (Grover & Pea, 2013; Pea & Kurland, 1984). The unplugged mode fosters the understanding that CT processes do not occur only within digital contexts, but have a wider application in real-life problem-solving.

The design and development of the game followed an iterative user-centred process (Fullerton, 2008). More specifically, we tested the first design ideas of the game with a custom-made life-size game as a pilot educational intervention with primary school children (Tsarava, 2016). Later on, we developed and tested the usability of an early prototype with primary school students during a short workshop session. After integrating feedback from both previous stages, we continued with the examination of users' game experience quantitatively with an adult population to ensure the games' appropriateness for children before evaluating the game with the target age group (Tsarava et al., 2018). Feedback from this study was integrated again and resulted in the latest version of the games. The final version of the games was evaluated for its game experience in the target age group. The results of this evaluation are reported in the current article.

2 Games Description

Crabs & Turtles (Tsarava et al., 2018) consists of three different games: i. *The Treasure Hunt*, ii. *The Race*, and iii. *Patterns*. The games are designed for children at primary school level, focusing specifically on 3rd and 4th graders. They are intended to be used as integrated educational interventions in the classroom. Teachers play a central role in their implementation by acting as the game master in all three games, which can be played independently from each other and at any order of preference. The games aim at introducing and training processes related to CT, like abstraction, algorithms, decomposition, evaluation, and patterns. In particular, they focus on mathematical (i.e., addition, multiplication, subtraction, and angular degrees) and coding (i.e., conditionals, constants and variables, events, loops, operators, and sequences) concepts related to those processes.

Figure 1. The Treasure Hunt game/grid board: 1. Sequence of commands created by the players, 2. Pawn, 3. Treasure collection point, 4. Pawn with food treasure items and badges that are collected by the players.

The Treasure Hunt (see Figure 1) is the first game of *Crabs & Turtles.* Players have to strategically move the pawn in teams of two on a grid board to collect food treasure items for their pawns (either crabs or turtles). To do so, teams of two have to efficiently build sequences of instructions, consisting of specific card commands to move their pawns across the board to gather treasures. They also need to obey specific rules and restrictions on movements indicated by the environment. For example, crab and turtle pawns can move only across specific coloured tiles on the grid, water or stone, and grass or stone, respectively. The main learning objective of the game is the general introduction to algorithmic thinking and sequential problem solving, as well as the consideration of restrictions and the use of simple conditional orders. Coding concepts explicitly addressed in this game are sequences and loops. For successful application of coding concepts, players are awarded badges during the game (e.g., loop badge, sequence badge, etc.). Along with coding concepts, students get familiar with handling angular degrees in spatial orientation. The winner of the game is the team that first collects a specific number of food treasure items.

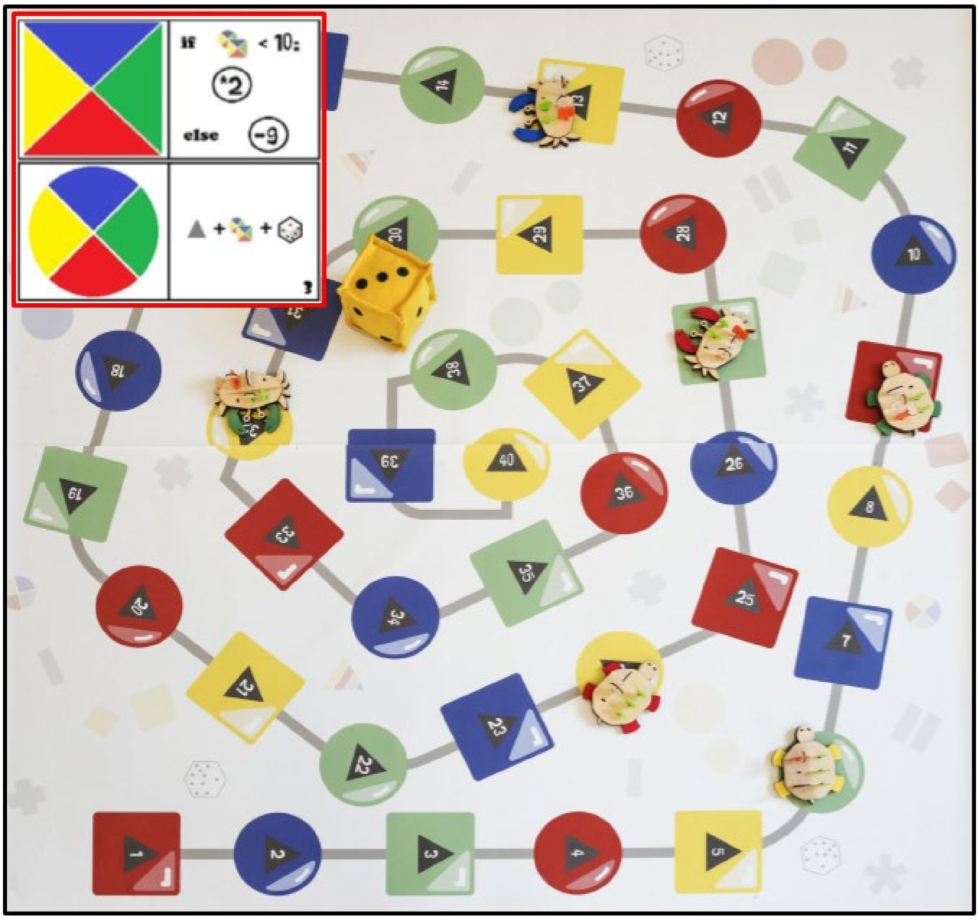

Figure 2. The Race game board (Inner upper panel: example of game cards).

The Race (see Figure 2) is the second game of *Crabs & Turtles.* In this game, players in teams of two have to reach the end of the game board by solving math/related riddles and handle the changing characteristics of variables (e.g., in-/decreasing of values). This game specifically focuses on coding concepts related to mathematics. In particular, coding skills explicitly addressed in this game are constants and variables, conditionals, events, and operators. During the game, players are awarded badges related to their achievements like variable badge, addition badge, etc. Mathematical abilities trained in the game relate to addition, subtraction, and multiplication. Consequently, the riddles of the game consist of equations related to mathematical operations and variables. The winner of the game is the team of two that first reaches the end of the race in the centre of the game board.

Patterns (see Figure 3) is the third game of *Crabs & Turtles.* In this game, children play individually, trying to match as fast as possible two types of cards based on visual patterns depicted on them. In order to do so, they have to read colour codes, recognize patterns, and follow specific restrictions. The colour codes consist of colours, a shape, and an arrow that indicates the order of reading the colour code (see Figure 3, left). The patterns consist of colourful shapes, which are depictions of a star, a square, a circle, and a triangle (see Figure 3, right). The order of the shapes, as well as their colour, is different on each card, matching in this way only one specific colour code. The main learning objective of the game is the introduction to the concept of patterns by identifying colour and shape patterns. The winner of the game is the player that succeeds in collecting the most cards.

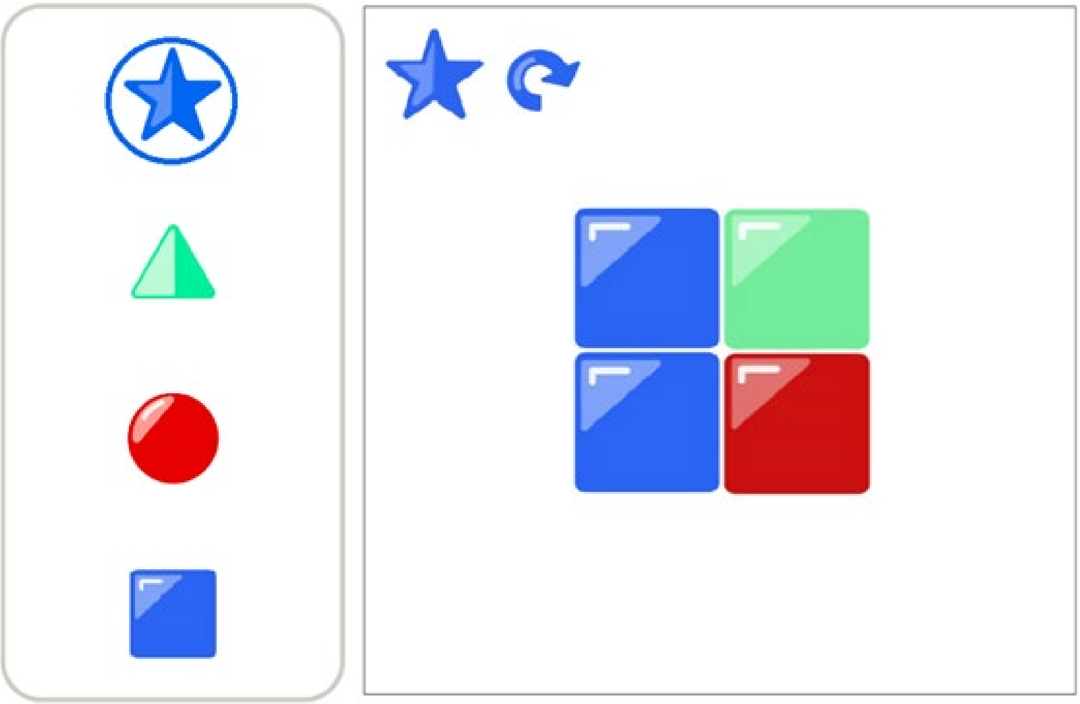

Figure 3. Patterns card pairing example (Left: a card depicting a colourful pattern; Right: a colour code matching the pattern card on the left).

3 Evaluation

After a successful 2-phase user test evaluation procedure with adult participants (Tsarava et al., 2018), we moved on to evaluating the games with primary school children – the actual target group of the games. In our 45-minute gaming sessions, the main focus was on assessing game experience quantitatively to identify potential dysfunctionalities during game-play, which then can be addressed before integrating the games into our CT course and evaluating their educational potential. To validate the design approach, participants consisted of different grades of primary school. Instructors and game masters in those sessions were the creators of the games.

3.1 Participants

We collected data from 79 primary school students aged between 8 and 12 years of age from 6 different schools in Greece and Germany. Due to missing data on more than 10% of the items, we excluded data of 9 participants from further analysis. For another 4 participants who completed more than one game, we had to exclude some of their questionnaires for specific games because responses were missing due to local organizational issues. Missing values for fewer items in the questionnaires were replaced by the mean score for the respective item computed from other participants. As such, data of a final sample of 70 participants was considered in the analyses (age in years: *mean* = 9.44, *SD* = 0.845; male: 42, female: 20, not indicated: 8).

3.2 Procedure and materials

In separate teaching sessions, we evaluated the game experience of primary school students. Most of the participants played all 3 games of *Crabs & Turtles*. Before participants started playing each of the games, we provided oral and visual instructions. After playing each game, participants were asked to complete the Game Experience Questionnaire (henceforth GEQ; Poels et al., 2007). We used a translated version of the Core (33 items) module in Greek and German to assess the overall game experience. The Core module consists of seven subscales addressing i. Immersion, ii. Flow, iii. Competence, iv. Positive Affect, v. Negative Affect, vi. Tension, and vii. Challenge. For each subscale, we used the average scores of the respective items as the dependent variable in our analyses. Each item had to be responded to on a 5-point Likert-scale (1 = not at all; 2 = slightly; 3 = moderately; 4 = fairly; 5 = extremely). For example, the fourth item of the Core module reads as follows: *"I felt happy",* and participants had to rate their experience of content on the aforementioned Likert scale by crossing an answer from 1 to 5 (e.g., crossing 4 would mean *"I felt fairly happy"*).

Furthermore, we used 4 additional items to further evaluate the overall game experience, which also employed a 5-point Likert-scale: *Q1. I would explain my experience as playing*; *Q2*. *I would explain my experience as learning* (Q1 & Q2: 1 = not at all; 2 = not really; 3 = undecided; 4 = somewhat; 5 = very much); *Q3*. *I would recommend the games to a friend*; *Q4. I would like to play the games again in the future* (Q3 & Q4: 1 = not at all; 2 = not really; 3 = undecided; 4 = likely; 5 = very likely). We added these 4 items to the questionnaire with the intention to measure the experience of the game as learning and/or playing because the GEQ aims at evaluating game experience more broadly and not game experience for educational games in particular.

Finally, to evaluate specific design elements of *The Treasure Hunt* and *The Race*, such as boards, cards, game pieces, inventory items, and rules, 5 more items (e.g., *Q: How much did you like the inventory items?*), again using a 5-point Likert-scale (1 = not at all; 2 = slightly; 3 = moderately; 4 = fairly; 5 = extremely), were used. For *Patterns,* only cards and rules were evaluated.

3.3 Results

The analyses of the questionnaires were conducted for each game separately. The current results are presented in the following three sections. We used a conservative approach of analyzing each subscale of the GEQ by conducting one-sample *t*-test comparing means of subscale ratings to the middle value of the scale (3 = mediocre) of the 5-point Likert scale. Internal consistency (Cronbach's Alpha) of the GEQ, as reported by Poels et al. (2007), is presented in Table 1 (column α*). In addition, Cronbach's alpha, as obtained in the current sample, is also reported in Table 1 (column α). The observed internal consistency indicated acceptable reliability for most subscales with $\alpha > .70$. However, this was not the case for subscales *Tension/Annoyance* (for games 2 and 3), *Challenge* (for games 1, 2, and 3), and *Negative Affect* (for game 1). For the analyses of overall game experience and the specific design elements, we again ran *t*-tests against the middle of the respective scale. Descriptive results and inferential statistics for the GEQ subscales are summarized in Table 1 and Figure 4.

3.3.1 The Treasure Hunt

Game Experience. Participants rated this game significantly above mediocre on the subscales *Competence*, *Sensory & Imaginative Immersion*, and *Positive Affect*. In contrast, ratings were significantly below mediocre for the subscales *Tension/Annoyance*, *Challenge*, and *Negative Affect*. We did not find a significant difference from mediocre for *Flow* (see Table 1).

Overall Experience. Participants experienced *The Treasure Hunt* somewhat as playing (Q1: *mean* = 4.22, *SD* = 1.10; $t(17) = 6.29$, $p < 0.001$) as reflected by ratings significantly above mediocre and not so much as a learning activity (Q2: *mean* = 2.98, *SD* = 1.51; $t(17) = -.09$, $p = 0.929$). Additionally, participants reported that they would likely recommend the game to a friend (Q3: *mean* = 3.74, *SD* = 1.37; $t(17) = 3.05$, $p = 0.005$), and would likely play the game again in the future (Q4: *mean* = 4.29, *SD* = 1.08; $t(17) = 6.71$, $p < 0.001$), as indicated by ratings significantly above mediocre.

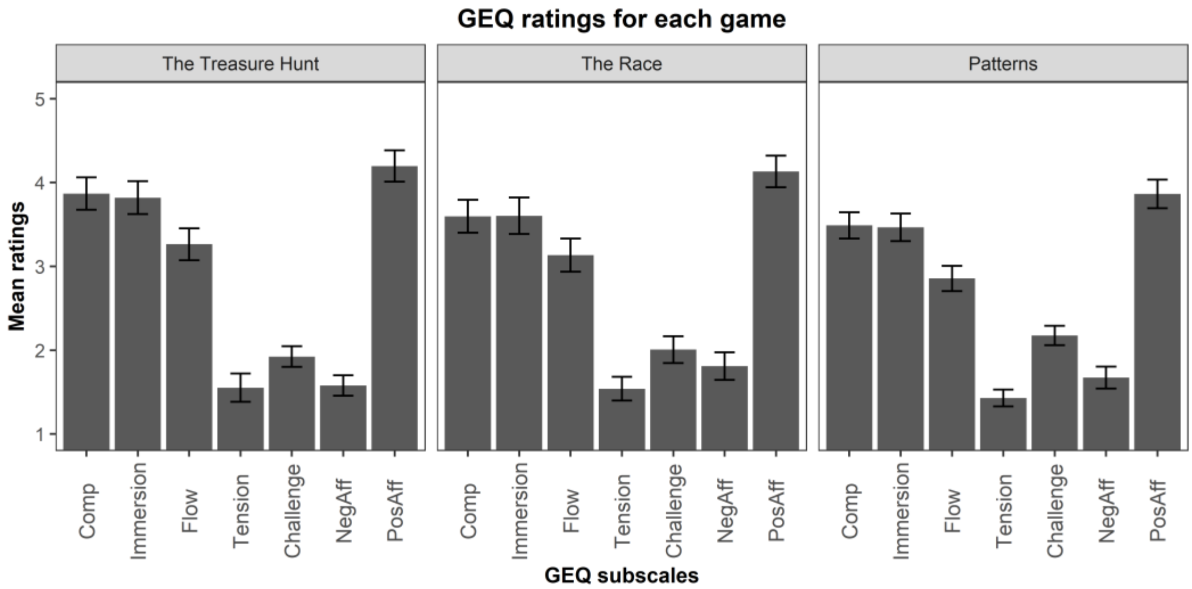

Figure 4. Students' ratings of GEQ subscales for each of the three games. On the y-axes, mean ratings of each subscale of the GEQ is represented. The y-axes refer to each of the subscales of the GEQ (Comp = Competence; Immersion = Sensory & Imaginative Immersion; Flow = Flow; Tension = Tension/Annoyance; Challenge = Challenge; NegAff = Negative Affect; PosAff = Positive Affect). Error bars depict 1 standard error of the mean.

Design Elements' Evaluation. The design elements of *The Treasure Hunt* scored a mean of 4.13 (*SD* = 1.02) on the 5-point Likert scale. More specifically, users rated all five design elements of *The Treasure Hunt* (Board: *mean =*4.09*, SD =1.17*, $t(17) = 5.27$, $p < 0.001$; Cards: *mean =*

3.88*, SD* = 1.24, *t(17)* = 4.00, *p* < 0.001; Game Pieces: *mean* = 4.50, *SD* = .84, *t(17)* = 10.07, *p* < 0.001; Inventory items: *mean* = 4.16, *SD* = 1.17, *t(17)* = 5.61, *p* < 0.001 and Rules: *mean* = 4.02, *SD* = 1.35, *t(17)* = 4.27, *p* < 0.001) significantly above mediocre.

3.3.2 The Race

Game Experience. Participants' ratings for this game were significantly above mediocre for the *Competence*, *Sensory & Imaginative Immersion*, and *Positive Affect* subscales of the GEQ Core module. In contrast, participants rated the game significantly below mediocre on the subscales *Tension/Annoyance*, *Challenge*, and *Negative Affect*. Also, we did not find a significant difference to mediocre for the *Flow* subscale.

Overall Experience. For *The Race* ratings significantly above mediocre indicated that participants rated their game experience somewhat as playing (Q1: *mean* = 4.16, *SD* = 1.03; *t(16)* = 5.64, *p* < 0.001) and not as a learning activity for which there was no significant difference from mediocre (Q2: *mean* = 3.21, *SD* = 1.41; *t(17)* = 0.74, *p* = 0.467). Furthermore, ratings significantly above mediocre reflected that they would likely recommend the game to a friend (Q3: *mean* = 3.91, *SD* = 1.12; *t(16)* = 4.09, *p* < 0.001), and also would likely play it again in the future (Q4: *mean* = 4.09, *SD* = 1.15; *t(16)* = 4.72, *p* < 0.001).

Design Elements' Evaluation. Overall, all five design elements of *The* Race were positively rated scoring a mean of 4.02 (*SD* = 1.01). More specifically, participants liked all five design elements (Board: *mean* =4.08*, SD =1.08*, *t(16)* = 5.03, *p* < 0.001; Cards: *mean* = 3.75*, SD* = 1.20, *t(16)* = 3.13, *p* = 0.005; Game pieces: *mean* = 4.42, *SD* = 1.08, *t(16)* = 6.58, *p* < 0.001; Inventory items: *mean* = 3.91, *SD* = 1.22, *t(16)* = 3.72, *p* = 0.001 and Rules: *mean* = 3.96, *SD* = 1.27, *t(16)* = 3.76, *p* = 0.001) as reflected by ratings significantly above mediocre.

3.3.3 Patterns

Game Experience. Similarly to the results of the other two games, participants' ratings for *Patterns* were significantly above mediocre for the *Competence*, *Sensory & Imaginative Immersion* and *Positive affect* subscales of the GEQ Core module. Again, ratings for the *Tension/Annoyance*, *Challenge*, and *Negative Affect* subscales were significantly below mediocre. Also, we did not find a significant difference to mediocre for the subscale *Flow*.

Overall Experience. Participants perceived the *Patterns* game somewhat as a playing experience (Q1: *mean* = 3.93, *SD* = 1.24; $t(15) = 4.92$, $p < 0.001$) which was again reflected by ratings above mediocre and with a marginally significant score as a learning activity as well (Q3: *mean* = 3.42, *SD* = 1.43; $t(17) = 1.94$, $p = 0.059$). Moreover, according to ratings above mediocre, participants reported that they would likely recommend it to a friend (Q3: *mean* = 3.74, *SD* = 1.29; $t(15) = 3.75$, $p = 0.001$), and also play the game again (Q4: *mean* = 4.10, *SD* = 1.30; $t(15) = 5.58$, $p < 0.001$).

Table 1. Mean scores for the Core module of GEQ at phase 2, per game-based activity.

The Treasure Hunt							
GEQ modules	***mean***	***SD***	***df***	***t***	***p***	***α***	***α****
Core							
Competence	3.86	1.09	31	4.48	0.000	.923	.826
Sensory & Imaginative Immersion	3.82	1.11	31	4.17	0.000	.910	.891
Flow	3.26	1.07	31	1.39	0.174	.813	.866
Tension/Annoyance	1.55	.96	31	-8.57	0.000	.890	.811
Challenge	1.92	.70	31	-8.76	0.000	.560	.745
Negative Affect	1.58	.69	31	-11.69	0.000	.640	.712
Positive Affect	4.20	1.06	31	6.38	0.000	.956	.797
The Race							
GEQ modules	***mean***	***SD***	***df***	***t***	***p***	***α***	***α****
Core							
Competence	3.60	.98	24	3.05	0.006	.799	.826
Sensory & Imaginative Immersion	3.60	1.09	24	2.77	0.011	.876	.891
Flow	3.13	.99	24	.676	0.506	.759	.866
Tension/Annoyance	1.54	.71	24	-10.34	0.000	.546	.811
Challenge	2.00	.80	24	-6.20	0.000	.673	.745
Negative Affect	1.81	.82	24	-7.28	0.000	.745	.712
Positive Affect	4.13	.95	24	5.98	0.000	.938	.797
Patterns							
GEQ modules	***mean***	***SD***	***df***	***t***	***p***	***α***	***α****
Core							
Competence	3.49	1.03	42	3.12	0.003	.865	.826
Sensory & Imaginative Immersion	3.46	1.08	42	2.82	0.007	.871	.891
Flow	2.86	.99	42	-0.95	0.345	.771	.866
Tension/Annoyance	1.43	.66	42	-15.72	0.000	.625	.811
Challenge	2.18	.75	42	-7.21	0.000	.592	.745
Negative Affect	1.67	.87	42	-10.06	0.000	.752	.712
Positive Affect	3.86	1.12	42	5.07	0.000	.928	.797

Design Elements' Evaluation. The design elements in *Patterns* scored a mean of 3.98 (*SD* = 1.01) on the 5-point Likert scale. The two design elements in the questionnaire for *Patterns* scored positively as indicated by ratings significantly above mediocre (Cards: *mean* = 4.00*, SD* = 1.05, *t(15)* = 6.27, *p* < 0.001; and Rules: *mean = 3.95*, *SD* = 1.09, *t(15)* = 5.72, *p* < 0.001).

4 Discussion and Future Work

The present study aimed at evaluating the game experience of primary school students in the three games of *Crabs & Turtles* and thus complements a previous evaluation of game experience in adults (Tsarava et al., 2018). After evaluating the game experience in adults and gathering overall positive results and valuable feedback, we completed the design of the final prototype for our games and tested user experience in the actual target group. We play-tested the games and collected data using the GEQ. Quantitative analyses on the game users' experience provided promising results regarding the validity of our approach.

Student participants rated their game experience after playing each game. Results indicated an overall positive reception of the games. In particular, students reported feeling competent and immersed while playing all three games, as well as experiencing positive affect. On the other hand, the overall challenge was rated low. Importantly, tension and negative affect ratings were also low for all three games. In addition, all three games of *Crabs & Turtles* were experienced as a playing activity, and students would likely be willing to play all three of them again and recommend them to their friends. Additionally, evaluation of the quality of design elements for each game was rated highly positive. In summary, this indicates that we managed to implement CT concepts into three gaming activities while achieving an overall positive game experience in children. However, the actual educational value of each of the games needs to be investigated comprehensively and evaluated empirically in separate studies, which, in fact, are currently being conducted in Germany.

The main aim of this study was the quantitative evaluation of primary school students' game experience in *Crabs & Turtles* to extend a previous evaluation in adults (Tsarava et al., 2018). The overall positive evaluation of game experience replicated in the target group now allows for a comprehensive evaluation of cognitive and educational benefits when playing the games. Although the overall results were positive, the relatively low scores in challenge and flow in all three games may not be optimal. Therefore, we plan to provide a set of game instructions with multiple adaptations. For example, we will facilitate selection of difficulty levels based on the number of players, so that the game becomes adaptive to classroom conditions (e.g., few or many students) and to students' game understanding (e.g., in case the game is understood well and gameplay seems easy, rules could become gradually more challenging while playing). We also plan to adapt a challenging game mechanic in *The Race* that will foster competition between the teams at every round of the game by allowing all the teams to solve the riddle as fast as possible.

Future studies are planned with primary school students of 3rd and 4th grade to evaluate the learning outcomes of the three games and their educational effectiveness on training CT-related skills. The three games, as part of a structured curriculum dedicated to training CT (Tsarava et al., 2017), will be evaluated through a pre-/post-test study design using a randomized field trial with a control group in 20 Hector Children's Academies in Baden-Wuerttemberg, Germany. Moreover, this forthcoming evaluation will aim at investigating cognitive abilities underlying CT and possible transfer effects of the course, using standardized cognitive tests to allow a diverse approach and definition of CT. Finally, we aim at developing digital versions of our board games to allow for individual dynamic adaptation of – for instance – the difficulty of the games.

References

Astrachan, O., & Briggs, A. (2012). *The CS Principles Project*. *3*(2), 38–42.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(August), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. *Annual American Educational Research Association Meeting, Vancouver, BC, Canada*, 1–25. http://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf

Brown, N. C. C., Sentance, S. U. E., Crick, T. O. M., & Humphreys, S. (2014). Restart: The Resurgence of Computer Science in UK Schools. *ACM Transactions on Computing Education*, *14*(2), 1–22. https://doi.org/10.1145/2602484

Code.org. (n.d.). https://code.org/

Echeverría, A., García-Campo, C., Nussbaum, M., Gil, F., Villalta, M., Améstica, M., & Echeverría, S. (2011). A framework for the design and integration of collaborative classroom games. *Computers and Education*, *57*(1), 1127–1136. https://doi.org/10.1016/j.compedu.2010.12.010

European Coding Initiative. (n.d.). http://www.allyouneediscode.eu/

European School Network. (n.d.). https://www.esnetwork.eu/

Fullerton, T. (2008). Game Design Workshop: A Playcentric Approach to Creating Innovative Games. In *Technology*. Elsevier Inc. https://doi.org/10.1007/s13398-014-0173-7.2

Grover, S., & Pea, R. (2013). Computational Thinking in K-12: A Review of the State of the Field. *Educational Researcher*, *42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

Haddad, R. J., & Kalaani, Y. (2015). *Can Computational Thinking Predict Academic Performance ?* 225–229.

National Science Foundation. (n.d.). https://www.nsf.gov/

Papert, S. (1994). Philosophy and the Computer. *Philosophical Books*, *35*(1), 39–41. https://doi.org/10.1111/j.1468-0149.1994.tb02396.x

Papert, S., & Solomon, C. (1971). *Twenty Things To Do With a Computer*.

Pea, R. D., & Kurland, D. M. (1984). On the cognitive effects of learning computer programming. *New Ideas in Psychology*, *2*(2), 137–168. https://doi.org/10.1016/0732-118X(84)90018-7

Poels, K., de Kort, Y., & Ijsselsteijn, W. (2007). *FUGA - The fun of gaming: Measuring the human experience of media enjoyment. Deliverable D3.3: Game Experience Questionnaire*.

Tsarava, K. (2016). *Programming in Greek with Python* [Aristotle Univerity of Thessaloniki]. http://ikee.lib.auth.gr/record/284234/files/GRI-2016-17183.PDF

Tsarava, K., Moeller, K., & Ninaus, M. (2018). Training Computational Thinking through board games: The case of Crabs & Turtles. *International Journal of Serious Games*, *5*(2), 25–44. https://doi.org/10.17083/ijsg.v5i2.248

Tsarava, K., Moeller, K., Pinkwart, N., Butz, M. V., Trautwein, U., & Ninaus, M. (2017). Training computational thinking: Game-based unplugged and plugged-in activities in primary school.

Proceedings of the 11th European Conference on Games Based Learning, ECGBL 2017, *October*, 687–695.

Tuomi, P., Multisilta, J., Saarikoski, P., & Suominen, J. (2018). Coding skills as a success factor for a society. *Education and Information Technologies*, *23*(1), 419–434. https://doi.org/10.1007/s10639-017-9611-4

Wang, P. S. (2015). *From Computing to Computational Thinking* (1st ed.). Chapman and Hall/CRC.

Wing, J. M. (2006). Computational Thinking. *Theoretical Computer Science*, *49*(3), 33–35. https://doi.org/https://www.cs.cmu.edu/~15110-s13/Wing06-ct.pdf

Wing, J. M. (2010). Computational Thinking: What and Why? *Thelink - The Magaizne of the Varnegie Mellon University School of Computer Science*, *March 2006*, 1–6. http://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational Thinking in Elementary and Secondary Teacher Education. *ACM Transactions on Computing Education*, *14*(1), 1–16. https://doi.org/10.1145/2576872

7 Computational Thinking: Cognitive Definition & Assessment

In this chapter, the following articles and manuscripts are attached:

- Study 4: **Cognitive Correlates of Computational Thinking: Evaluation of a Blended Unplugged/Plugged-In Course.**
- Study 5: **A Cognitive Approach to Defining and Assessing Computational Thinking: An Empirical Study in Primary School** (under review).
- Study 6: **Evaluation of a Computational Thinking Intervention for Elementary School Children: A Randomized Controlled Field Trial** (in preparation).

Cognitive Correlates of Computational Thinking: Evaluation of a Blended Unplugged/Plugged-In Course

Katerina Tsarava, Luzia Leifheit, Manuel Ninaus, Marcos Román-González, Martin V. Butz, Jessika Golle, Ulrich Trautwein, Korbinian Moeller

Abstract

Coding as a practical skill and computational thinking (CT) as a cognitive ability have become an important topic in education and research. It has been suggested that CT, as an early predictor of academic success, should be introduced and fostered early in education. However, there is no consensus on the underlying cognitive correlates of CT in young elementary school children. Therefore, the present work aimed at (i) assessing CT and investigating its associations to established cognitive abilities, and (ii) evaluating a newly developed CT course for elementary school children.

As such, 31 7-10-year-old children took part in 10 lessons of a structured CT course. The course aimed at introducing and fostering CT concepts in both unplugged and plugged-in ways, incorporating life-size board games, Scratch, Scratch for Arduino, and Open Roberta programming environments. In a pre-/post-test design, we assessed several cognitive abilities using standardized tests on non-verbal visuospatial and verbal reasoning abilities, numeracy, as well as short-term memory and measured CT using an adapted version of the only existing validated test *CTt*, to accommodate it to the younger sample.

We identified significant associations between CT and non-verbal visuospatial reasoning, as well as different aspects of numeracy (e.g., fact retrieval and problem completion). In line with recent theoretical accounts and empirical investigations for other age groups, these findings specify the underlying cognitive mechanism of CT in elementary school. Moreover, our results indicated that students were able to specifically improve their CT abilities through the course, as assessed by the adapted version of the *CTt*.

Keywords: computational thinking, computational thinking assessment, computational thinking curriculum, cognitive skills

1 Introduction

1.1 Definition of CT

Computational thinking (CT) has been coined a crucial 21st-century skill, along with collaboration, communication, digital literacy, citizenship, problem-solving, critical thinking, creativity, and productivity (Bocconi et al., 2016; Voogt et al., 2013). As such, these skills have been suggested to be important for children to acquire and develop early on in education to be prepared for the demands of the present and future digital world (Settle et al., 2013; Yadav et al., 2014, 2011). The first conceptualized introduction of the term CT by Wing (Wing, 2006a) described it as an applicable attitude and skill set for everyone and not only for programmers or computer scientists. Due to the term attracting considerable interest in research and education alike, there have been numerous attempts to further defining and conceptualizing CT (e.g., Barr et al., 2011; Grover & Pea, 2013).

A recent review on CT in educational contexts (Shute et al., 2017) describes CT as a "conceptual foundation required to solve problems effectively and efficiently (i.e., algorithmically, with or without the assistance of computers) with solutions that are reusable in different contexts". This holistic definition characterizes CT as a cognitive ability rather than just being a practical skill in a specific context and hence emphasizes the broad applicability of CT. This is particularly important, as there are common misconceptions between the terms and definitions of CT, computer science, programming, and coding. As part of computer science, coding and programming usually refer to the somewhat more practical and less theoretical skills of computer science, referring to writing computer code and building computer programs, respectively. In contrast, CT is considered as a broader cognitive concept, which is associated with computer science, but which is also applied more broadly in various other - not essentially computerized - domains (Armoni, 2016). While it is assumed that basic CT competencies are required for acquiring more practical coding and programming skills, CT generally reflects a broader cognitive skill that is crucial for computational literacy (Balanskat & Engelhardt, 2015; García-Peñalvo et al., 2016). Moreover, recently, the definition of CT was extended by Resnick (2017), who used the term computational fluency to describe CT not only as an understanding of computational concepts and problem-solving strategies but also as a creative ability for self-expression through means of digital technologies.

1.2 CT curriculum and assessment

Although there are inconsistencies in the definition and conceptualization of CT, its value for education is broadly accepted (V. Barr & Stephenson, 2011; Qualls & Sherrell, 2010). Consequently, during the past decade, many governments (e.g., for England, see Williamson, 2015) and educational institutions all over the world adopted and integrated CT in their national curricula for elementary (e.g., Bers et al., 2014), secondary (e.g., Settle et al., 2012), and post-secondary education (e.g., Dierbach et al., 2011). This integration was based on different suggested frameworks for designing CT curricula (Curzon et al., 2014; Perković et al., 2010) aiming at fostering CT as a broader cognitive ability applicable in different contexts. The continued interest in CT resulted in numerous newly designed educational materials to be used for fostering CT. However, at the same time, tools or instruments measuring CT are required to assess the outcomes of these educational materials.

The quantitative assessment of CT may be contradictory to the initial anthropological views of constructivism, which proposed the idea that learning is the active construction of meaning rather than the simple internalization of knowledge (Hein, 1991) and can therefore not adequately be measured by assessing individual, prespecified learning outcomes. However, educators should monitor learning outcomes to assess the quality of the used educational material (Hickmott & Prieto-Rodriguez, 2018). The assessment of CT has so far been attempted in different ways (e.g., Hubwieser & Mühling, 2014; Moreno-León et al., 2015; Román-González et al., 2017; Weintrop et al., 2014).

Brennan and Resnick (Brennan & Resnick, 2012b) suggested assessments based on project portfolio analyses, artefact-based interviews, and design scenarios in the Scratch programming environment. Assessment components of their approach were extended and integrated into the Progression of Early Computational Thinking (PECT) model (Seiter & Foreman, 2013), which is comprised of evidence variables (e.g., conditionals, operators, boolean expressions), design pattern variables (e.g., animated looks, animated motion, user interaction), and CT concepts (e.g., procedures and algorithms, problem decomposition, parallelization, and synchronization), based on which Scratch projects are assessed.

Weintrop et al. (2014) proposed a digital interactive assessment of different items designed to measure students' CT ability within STEM disciplines. Dr. Scratch (Moreno-León et al., 2015) is a formative assessment tool also based on the analysis of CT concepts identified in Scratch projects, implemented in an automatic way. The Fairy assessment (L. Werner, Denner, & Campe, 2012) measures students CT abilities while engaging students in the game programming environment of Alice. Koh et al. (2014) proposed a real-time assessment tool that measures students CT performance while they engage the game programming environments in Agentsheet and Agentcubes activities. Most of these assessment methods are linked to certain digital programming environments, such as Scratch.

An early psychometric approach for assessing CT, detached from a specific digital environment, is the work of Mühling et al. (2015), who created and evaluated an assessment tool that focuses on control flow concepts in programming algorithms. Building upon previous work, Weintrop and Wilensky (2015) developed the Commutative Assessment, which assesses with 28 items different CT concepts, such as conditional logic, loops, functions, algorithms, variables, and comprehension in two different modalities, text- and block-based. Ambrósio et al. (2015) created a computerized cognitive assessment for CT, which is implemented as a tablet-based test. Another psychometric approach assesses CT in robotics programming and everyday reasoning with 5th-grade students using 15 multiple choice and 8 open-ended questions in pen-and-paper form (Chen et al., 2017). Those 23 items were developed based on 6 CT components, namely syntax, data, algorithms, representation of problems and solutions, as well as generation of efficient and effective solutions. Furthermore, in an attempt to assess CT in preschoolers, the Coding Development (CODE) Test 3-6 (Marinus et al., 2018) was developed.

From this brief review, it becomes evident that many different instruments seem to be available for assessing CT. However, most of them are not evaluated and validated in a large sample. One clear exception to this is the Computational Thinking Test (*CTt*), which has been validated with over 1200 children (Román-González et al., 2017). The *CTt* comprises 28 items, addressing one or more out of 7 different CT concepts in multiple-choice tasks (Román-González, 2015b). These are presented primarily utilizing a Pac-Man maze or a painting canvas and thus do not specifically refer to a programming language such as Scratch, for instance. Studies on *CTt* so far provide empirical evidence on its content validity (Román-González, 2015b), criterion validity (Román-González et al., 2018b; Román-González et al., 2017), convergent validity (Román-González, Moreno-León, et al., 2017), predictive validity (Román-

González et al., 2018a), and cross-cultural validity (Wiebe et al., 2019). *CTt* was used in the current study to investigate the association between CT and other cognitive abilities to understand its underlying processes better.

1.3 Cognitive correlates of CT

Already back in 1977, Brooks (1999) pointed out the necessity of more theory-driven studies focusing on specifying the underlying cognitive processes of programming. The author based his theory of programming on the problem-solving theory by Newell and Simon (Simon & Newell, 2006). Accordingly, he characterized the underlying cognitive processes of programming as understanding a written task and finding appropriate methods and strategies for solving it. In this context, a meta-analysis of 65 studies on the cognitive outcomes of programming (Liao & Bright, 1991) revealed that computer programming could have positive effects on a range of cognitive abilities, such as reasoning skills, logical thinking, planning skills, and general problem-solving skills. A more recent meta-analysis of 105 studies on the effect of computer programming on improving other cognitive skills (Scherer et al., 2019) provided further evidence for cognitive benefits in terms of near and far transfer effect. For instance, there was a positive transfer found in situations that required creative thinking, mathematical skills, and metacognition, followed by spatial skills and reasoning.

Recently, there have been quite a few studies focused on programming behaviours and on assessment of cognitive processes while coding in specific programming environments. However, most of them fall short of providing more detailed empirical insights into explicit cognitive correlates of CT. Nevertheless, the recent increase in educational materials and courses to foster CT also surged interest in examining specific cognitive processes underlying CT. While this is not only highly relevant for the proper assessment of existing courses that are supposed to teach and improve CT abilities, it also seems important for advancing the conceptualization and definition of CT.

However, empirical studies on the cognitive correlates of CT are still very sparse. Recent studies on the cognitive underpinnings of coding with young children of 5 to 6 years of age (Marinus et al., 2018) showed that cognitive compiling (the ability to formulate mental action plans in natural language) was a predictor of coding ability over and beyond influences of age and non-verbal intelligence. In another study, Román-González and colleagues (2017) not only validated the *CTt* but also examined associations between the *CTt* and other cognitive skills. For the participating students (10-16 years of age), there were weak to moderate correlations between CT and three of the four primary mental abilities assessed (i.e., reasoning, spatial, and verbal ability), whereas a high correlation was observed between CT and problem-solving abilities. However, no statistically significant correlation was observed between CT and numerical ability. These findings are consistent with recent theoretical proposals of Ambrósio et al. (2015) that associate CT with some of the core elements of the Cattell-Horn-Carroll (CHC) model of intelligence (McGrew, 2009), especially with respect to fluid reasoning (Gf), visual processing (Gv) and short-term memory (Gsm).

1.4 Aim of the study

The current study aimed at investigating whether similar associations between CT and cognitive abilities, as shown by Román-González et al. (2017), can also be found in elementary school students, for whom the cognitive correlates of CT have not been evaluated so far. Moreover, we asked the question of whether a course designed to improve CT abilities can improve both *CTt* scores as well as the scores in other previously correlated mental abilities. Therefore, the current study employed a pre-/post-test pre-experimental design assessing CT as well as relevant cognitive abilities before and after a newly developed course, which was designed to effectively foster CT. As a result, it was possible not only to analyze associations between CT and other cognitive abilities but also which of them would be affected by the teaching intervention targeting CT. Based on the most recent literature on CT, we had the following hypotheses:

(i) We hypothesize to observe same (or similar) correlation patterns between *CTt* scores and cognitive abilities in elementary school children when compared to those found in secondary school children, like reasoning, spatial and verbal abilities.

(ii) We expect to find children's *CTt* performance to be improved after participating in the targeted CT training course.

2 Methods

2.1 Participants

Participants were 31 (28 boys) 3rd and 4th graders enrolled in the extra-curricular programs of four Hector Children's Academies in Baden-Württemberg, Germany. Students' age varied between 7 and 10 years (mean = 9.47; SD = 0.76). All students decided voluntarily which extra-curricular course or courses they wanted to attend from a list of various offered courses in each academy. The students participating in the study attended a CT training course consisting of 10 90-minutes lessons. Informed consent was provided by parents before the start of the study.

2.2 Study design

Participants took part in a ten-week CT training course. One week before and one week after the training, different cognitive abilities, as well as CT, were assessed in two 2-hour long sessions each, using the same battery of tests. The study took place in four different local extra-curricular programs. The intervention was taught by one instructor in each class. Instructors of the course were one of the researchers at three of the classes and one qualified computer science teacher in one class. In the following, we first describe the general pedagogical principles utilized in the developed CT course before we briefly report the content of each lesson. The course consists of 10 lessons of 90 minutes each. It is a blended course incorporating both unplugged and plugged-in interventions and is structured in 4 sections: i. Unplugged introduction to CT concepts using life-size board-games (Tsarava et al., 2018), ii. Application of CT concepts in the Scratch (Lifelong Kindergarten Group - MIT Media lab) programming environment, iii. Transfer of CT concepts to the Scratch for Arduino (S4A; Citilab, 2015) programming environment, and iv. Robot simulation using CT concepts in the Open Roberta Lab (Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS) robot programming and simulation environment. We documented the theoretical and methodological approaches of the course, as well as its' educational teaching materials (e.g., lesson plans, step-by-step activities, game materials, short assessments, etc.) in a detailed course manual of 250 pages.

2.2.1 Conceptual teaching approach

The intervention followed a conceptual teaching approach closely linked to CT, aiming at imparting computer science knowledge and competencies, which are quickly learnable, well structured, and easily transferable to different domains. In order for children to internalize this knowledge and develop the competencies as sustainably as possible, the intervention's contents are introduced in a conceptual way. That is, students should be able to independently work on programming projects after completion of the current training as well as understand and approach computational problems. Consequently, the training was not on specific programming environments or technologies but aimed at fostering basic CT processes and coding concepts, such as sequences, loops, conditional branching, events, variables, and operators (as identified by Brennan & Resnick (2012).

The first introduction to CT in the course is done by the use of unplugged activities. Such an approach was proposed by Prottsman (Prottsman, 2014) to familiarize younger students with computational topics and problems. The aim was to these use unplugged exercises to facilitate the acquisition of abstract computational concepts, which students will implicitly encounter when they first begin to program (Prottsman, 2014) because the attainment of practical programming skills requires at least knowledge of basic computational concepts (Balanskat & Engelhardt, 2015; García-Peñalvo et al., 2016). However, the promotion of CT detached entirely from actual programming may lead to artificial and contextless learning scenarios. Therefore, it has been suggested to promote CT by teaching it in a conceptual way, applying it in practical programming, and transferring it to interdisciplinary contexts (Yadav et al., 2014).

2.2.2 Game-based learning

Learning activities in which learners build knowledge and competences through games or playful tasks are referred to as game-based learning activities. The use of game-based teaching and learning methods can improve learners' performance (e.g., Ninaus et al., 2015; for a meta-analysis, see Wouters et al., 2013) and increase their motivation and involvement in the learning activity (Hamari et al., 2014). Games and game-based activities are an increasingly important approach for cognitive training, learning, and education (Boyle et al., 2016b) as they motivate learners to interact actively with the learning environment (Plass et al., 2015).

Importantly, game-based learning also seems to be superior in terms of learning performance as compared to traditional teaching methods (Wouters et al., 2013).

Accordingly, the present CT training course utilizes life-size board and card games we developed specifically for the training. Moreover, previous studies indicated that 3rd and 4th grade students benefit particularly from unplugged game activities to introduce CT concepts as this format allows for a low threshold introduction facilitating understanding and application of these concepts (Leifheit et al., 2018). In particular, these games aim at introducing different computational concepts and increasing active involvement of the players and thus the learning content (for an overview, see Echeverría et al., 2011). Moreover, the games' board and the card games aim at making abstract coding concepts tangible, thus aiding the development of abstract and symbolic thinking through multimodal representations (Plass et al., 2015).

2.2.3 Embodied Learning

The theory of embodied cognition states that many aspects of human cognition are grounded in physical, interactive experiences of the learner with the environment (Barsalou, 2008). In concrete learning scenarios, this is the case when the body of the learner plays a significant causal role in the experienced situated dynamics (García-Peñalvo et al., 2016). Accordingly, learning activities in which learners acquire knowledge and competences through physical interaction or manipulation of physical objects are referred to as embodied learning.

Embodied teaching and learning methods can facilitate implicit learning (Barsalou, 2008) and can have a positive effect on comprehension as well as retention by spatially organizing and representing conceptual content (Noice & Noice, 2001).

The life-size board and card games used in the present course are intended to enable embodied experiences of the basic concepts and processes of CT (Barsalou, 2008), on the basis of which the development of conceptual, event-predictive abstractions can be supported in a maximally natural manner (Butz, 2016; Butz & Kutter, 2017). For that reason, the design of the game board grid takes into consideration the fee size of a young child in order to allow for active movement on it. In one of these board games, for example, the concept of a loop is introduced and practised by recognizing repetitions in sequences of movements of a game character and then reformulating them as loops of sequences of movement.

Furthermore, the open-hardware platform Arduino, as well as the programming language and environment S4A (Scratch for Arduino), was used in the course to enable embodied learning. By controlling the different sensors of the Arduino and programming the functionality of the electronic components, students experience that they can not only influence what happens on the screen before their eyes but can also interact with the physical world around them through the programs they write.

2.2.4 Outline of the CT training

During the course of the training, students gradually progressed from hands-on and non-digital unplugged tasks (e.g., board games, pen-and-paper exercises; see Figure 2) to increasingly more abstract and sophisticated plugged-in programming tasks (for the course overview, see Figure 1). In lessons 1, 2, and 3, students got to know basic CT concepts in an unplugged way by playing life-size board and card games (Tsarava et al., 2018; Tsarava, Moeller, et al., 2019). In lessons 2, 4, 5, and 6, students were introduced to the visual block-based programming language and environment Scratch. Hence, students got to apply the conceptual knowledge acquired while playing the board and card games to Scratch.

Tools	Unplugged life-size games & Scratch			Scratch			S4A		Open Roberta Lab	
Lesson	1	2	3	4	5	6	7	8	9	10
Mode Unplugged Plugged-in										

Figure 1. Course overview.

In lessons 7 and 8, when working with the Arduino hardware platform equipped with physical sensors, they experienced their program code engaging with the physical world around them.

In the last two lessons of the intervention, lessons 9 and 10, students exercised their CT abilities independently by programming a simulated robot in the interactive robot simulation environment Open Roberta Lab. This gave the students the opportunity to test the knowledge and skills they acquired throughout the training in a new context and to work more independently than in lessons 1 to 8. Open Roberta Lab is particularly suited for this type of work for several reasons: First, Open Roberta Lab provides an easy-to-learn block language. Thus, students can quickly find their way around the new programming language and environment. Second, the use of the predefined but easily accessible simulation environment ensured that students did not spend too much time on tinkering with the purely graphical elements of their programs, but actually had to program and think strategically in order to control the simulated robot. Finally, controlling the robot stimulates attempts on free problem formulation and solving. Specifically, there is no predetermined problem-solving approach, but students got to set their own goals, which could be achieved in an indefinite number of ways.

Each of the 10 lessons referred to a specific STEM discipline to which the respective applications and unplugged activities were closely related (for an overview of the connection between STEM disciplines and computational abilities, see D. Barr et al. (2011) and Sanders (2009). For example, students have to program simulations of a food chain or a water cycle in one of the lessons, addresses issues from biology or geography in another. This aimed at widening the students' perspective on CT and actual programming skills being valuable for solving problems in various contexts and not only for computer science.

Figure 2. An example of an unplugged activity of the life-size board games which introduces sequences, loops, simple conditionals, and events.

2.3 Instruments

To assess the association between CT and relevant other cognitive abilities, we used the following test battery in the pre- and post-test:

(i) For the assessment of arithmetic operations, five subtests of the Heidelberger Rechentest zur Erfasssung mathematischer Basiskompetenzen im Grundschulalter (HRT 1-4; Haffner et al., 2005) were used; namely, the subscales Writing Speed (30 seconds), Addition (2 minutes), Subtraction (2 minutes), Multiplication (fact retrieval; 2 minutes), and Problem Completion (2 minutes).

(ii) Non-verbal visuospatial reasoning was assessed using two subtests of the Culture Fair Intelligence Scale (CFT 20-R; Weiß, 2006), namely the Continuing Series (Subtest 1; 4 minutes) and the Matrices (Subtest 3; 3 minutes). Participants had 7 minutes to complete both subtests.

(iii) Verbal reasoning ability was assessed using the subtest V1 (Form A) of the Kognitiver Fähigkeitstest (KFT 4-12+R; Heller & Perleth, 2000), for which participants had 7 minutes to complete it.

(iv) CT abilities were assessed using 21 items of the Computational Thinking test (*CTt*; (Román-González et al., 2017) translated to German. Selected items were the 21 easiest items out of the total number of 28 items (Román-González et al., 2017). The respective original item numbers are as follows: 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 13, 14, 17, 18, 19, 20, 21, 24, 26, 27 and 28. This was done because of the considerable younger sample in the current study as compared to the original sample for which the test was validated by Roman-Gonzalez et al. (2017). The 7 different CT concepts addressed are i. basic directions and sequences, ii. loops implemented with repeat-times commands, iii. loops implemented with repeat-until commands, iv. simple conditionals implemented with if commands, v. complex conditionals, implemented with if/else commands, vi. while conditionals, and vii. simple functions. Participants had 20 minutes to complete the test.

2.4 Analysis procedure

The collected data were digitized, processed, and analyzed using IBM SPSS software. From each one of the aforementioned tests, we generated the following variables, coded as follows:

· WRSpre and WRSpost: number of correct responses in the Writing Speed subtest of the HRT 1-4, at pre- and post-test assessment.

· CALCpre and CALCpost: mean of correct responses in the Addition and Subtraction subtests.

· MULpre and MULpost: number of correct responses in the Multiplication subtest.

· COMpre and COMpost: number of correct responses in the Problem Completion subtest.

· CFTpre and CFTpost: mean of correct responses in the Continuing Series and Matrices subtests of the CFT.

· KFTpre and KFTpost: number of correct responses in the verbal reasoning subtest of the KFT.

· CTTpre and CTTpost: number of correct responses in the *CTt*.

To investigate associations between CT and other relevant cognitive abilities, we ran Pearson correlations between all variables at pre-test and post-test. To investigate whether CT performance at post-test was related to other relevant cognitive abilities at pre-test, we again ran Pearson correlations between the cognitive test scores at pre-test and the *CTt* values at post-test. Furthermore, to evaluate the developed CT course, paired t-tests were performed to compare pre- and post-test performance in *CTt* as well as the other assessed cognitive abilities to determine the specificity of the intervention. The effect sizes were calculated with Cohen's d for related samples. The reliability of the shortened version of *CTt* consisting of the 21 easiest items of the complete *CTt* version was measured with Cronbach's alpha.

3 Results

The Cronbach's alpha for the shortened version of *CTt* used for the current study was .766, as derived from the analysis on the pre-test performance scores of the test. This result indicates high internal consistency of the instrument, and we can assume that it is reliable for the age group that we administered it.

Results of the concurrent correlation analyses between CT and other cognitive abilities at pre-test are reported in Table 1. There was a moderately high positive correlation between performance in the *CTt* and performance at the HRT Multiplication subtest (r = .423; p < .022), as well as between *CTt* performance and HRT Problem Completion subtest (r = .398; p < .032).

These results suggest that CT abilities prior to any CT training were related to arithmetic abilities, like multiplication (fact retrieval) and problem completion.

Table 1: Performance correlations between cognitive tests and CTt at pre- and post-test.

Variable name	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) WRSpre														
(2) CALCpre	.378*													
(3) MULpre	.286	.812**												
(4) COMpre	.350	.881**	.702**											
(5) CFTpre	.159	.503**	.394*	.515**										
(6) KFTpre	.105	.171	.095	.291	.380*									
(7) CTTpre	.023	.347	.423*	.398*	.172	.108								
(8) WRSpost	.112	.231	.147	.130	-.032	-.200	.048							
(9) CALCpost	.399*	.771**	.701**	.755**	.796**	.398*	.410*	.139						
(10) MULpost	.265	.653**	.669**	.596**	.503*	.073	.169	.333	.620**					
(11) COMpost	.190	.622**	.495*	.757**	.580**	.183	.470*	.310	.649**	.694**				
(12) CFTpost	.169	.410*	.260	.382	.625**	.171	.272	.230	.504**	.529**	.572**			
(13) KFTpost	.338	.001	-.090	.072	.392*	.573**	.064	-.088	.279	-.035	.163	.264		
(14) CTTpost	.165	.119	.101	.190	.474*	.238	.217	.137	.411*	.324	.390*	.417*	.193	

*. The correlation is significant at the level of 0.05 (2-sided).

**. The correlation is significant at the level of 0.01 (2-sided).

The longitudinal correlation between CT and specific cognitive skills was evaluated using Pearson correlations between post-test *CTt* performance and the pre-test scores of the other cognitive abilities (see also lower left part of Table 1). Results suggested that only CFT seems to be predictive of CT performance at post-test assessment (r = .474; p < .014). This indicates high relevance of non-verbal visuospatial reasoning in CT training.

Finally, to investigate possible training effects due to our intervention we ran paired-samples t-tests between pre- and post-test scores for each cognitive test (see Table 2). A significant performance increase in *CTt* from pre- (M = 11.19; SD = 3.37) to post-test [M = 14.15; SD = 3.84] was detected [t(25) = 3.33, p = .003, d = .654] reflecting a medium effect size according to Cohen (1988). Moreover, there was a significant improvement from pre-test (M = 9.42; SD = 2.34) to post-test (M = 10.88; SD = 1.97) performance for non-verbal visuospatial reasoning [t(25) = 3.93, p = .001, d = .770] reflecting a rather large effect size.

Table 2: Paired-sample t-tests of cognitive tests and CTt at pre- and post-test.

	mean post-pre	std. deviation post-pre	t	degrees of freedom	sig. (2-tailed)
WRSpost - WRSpre	-2.00000	9.68676	-1.032	24	.312
CALCpost - CALCpre	-.38000	3.33004	-.571	24	.574
MULpost - MULpre	.04000	4.70354	.043	24	.966
COMpost - COMpre	.03846	4.88656	.040	25	.968
CFTpost - CFTpre	**1.46154**	1.89696	3.929	25	.001**
KFTpost - KFTpre	1.26923	3.31732	1.951	25	.062
CTTpost - CTTpre	**2.96154**	4.52973	3.334	25	.003**

4 Discussion

CT, as a 21st-century skill, driven by the progression of science and technology and the resulting demands for qualified individuals in STEM jobs, has lately resulted in the development of dedicated CT-centered and STEM-oriented educational materials. These materials are often addressed to a wide population of students, ranging from elementary school children to university level students. However, the evaluation of such dedicated materials aimed at fostering CT cannot be properly assessed without valid CT evaluation tools. Meanwhile, researchers and educators alike argue about a proper definition and conceptualization of CT. To advance the current state-of-the-art, we developed and ran a CT intervention course and assessed CT and a range of other cognitive abilities in elementary school children.

Our results showed significant and specific associations between CT and complex arithmetic abilities as well as non-verbal visuospatial reasoning abilities. Our observation of an association between CT and non-verbal visuospatial reasoning abilities replicates what was already found in (Román-González et al., 2017) for middle and secondary school, for elementary school children. This seems to indicate that non-verbal visuospatial reasoning abilities are moderate and consistent cognitive correlates of CT throughout different school levels. However, the associations between CT and complex arithmetic abilities that we found were unexpected based on previous results for middle and secondary school (Román-González et al., 2017). This is an interesting finding that could be explained from the fact that numerical abilities at an early age are needed as prerequisites for thinking computationally.

For example, to perform a simple sequence of commands, a young child should be able to count, or in order to apply a simple loop, one should be able to implement numerical operations like addition and multiplication. Nevertheless, once a certain threshold of numerical ability is exceeded at a certain level of school education, then it may not be equally determinant for higher levels of CT.

Furthermore, we observed no significant correlation between CT and verbal reasoning ability. Previous studies (Román-González et al., 2017) for middle and secondary school - although reporting a small correlation between verbal ability, also demonstrated that verbal ability per se did not predict CT in a regression analysis including other predictors. Synced with the current results, this indicates that verbal ability does not correlate with CT consistently across all school levels. Nevertheless, the results of the current pilot study need to be treated cautiously as these expected, and unexpected correlation patterns were not perfectly stable from pre- to post-test. This might be due to statistical power issues due to the small sample and needs to be investigated in the future with a larger sample size.

Moreover, as expected, the employed 10 lesson blended CT intervention improved children's CT abilities. Hence, these results are indicative of the effectiveness of the current CT course. Importantly, though, also verbal and non-verbal visuospatial reasoning skills increased through the CT training course. This is not only in line with the results of the correlational results at post-test, but these results also substantiate results from a previous study (Román-González et al., 2017), that showed moderately high correlations between CT and spatial abilities, reasoning, as well as problem-solving.

5 Limitations and Future Work

One limitation of the current work is the self-selection bias of our sample, which derives from the procedure of participating in our course. Students are not assigned to a specific course offered by their academies but are allowed to choose for themselves one or more courses they would like to participate in every semester, from a list of available courses. This means that our results could be generalized only to a similar population of students who voluntarily decide to attend a computational thinking course. Moreover, this self/selection procedure resulted in an unbalanced gender distribution with primarily boys attending the course.

Another limitation of this study is the existence of two different instructors teaching the four different classes. This might have an impact on the internal validity of our results. However, both instructors strictly followed a very detailed course manual and used the same teaching materials to ensure that the course was implemented as intended (i.e., with high implementation fidelity (Herbein et al., 2018). In addition, having two different instructors enabled us to get external feedback on the newly developed course manual, in order to further improve it for future instructors of the course.

Furthermore, the interpretation of the observed training effect is not straightforward because the current study did not include a control group. One might argue that just by completing the cognitive tests twice, performance may have improved. If this were the case though, we would expect rather similar improvements in all or at least the majority of the scales administered. However, improvements were only observed for CT - the primary aim of the current intervention - and non-verbal visuospatial reasoning abilities, which, as previous studies demonstrated, seem to be closely related to CT (e.g., Parkinson & Cutts, 2018; Román-González et al., 2017).

To the best of our knowledge, this is the first study that comprehensively investigates cognitive correlates of CT in the age group of elementary school children. Despite limitations of the current study in terms of statistical power due to the small sample size, our results are well in line with previous investigations of cognitive correlates of CT in other age groups (e.g., Parkinson & Cutts, 2018; Román-González et al., 2017). Additionally, our results also indicate that for elementary school children, arithmetical abilities also seem to be associated with CT. This expands previous findings for older children and may indicate that at an elementary level, there might be associations between first computational procedures in mathematics and CT as assessed by the *CTt*. However, one has to acknowledge that the *CTt* does not cover all aspects of CT deemed to be relevant in the literature, such as abstraction, decomposition, generalization, or algorithms directly. Nevertheless, the current study advances the ongoing debate on the conceptualization of CT by emphasizing the importance of numerical and spatial reasoning abilities for CT.

Future planned studies will extend the presented pilot study by implementing an experimental control group design to more than 15 different academies and up to 200 students, randomly assigned to groups. The course will be provided to the respective academies by instructed teachers. We expect this increase of sample size to avoid power issues, confirm and further substantiate current results.

ACKNOWLEDGEMENTS

We would like to thank for their enthusiastic involvement, the constant feedback, and support in conducting this study the following colleagues: Dominik-Johannes Krauß, Helmut Fritsch, and Jana Hoffstadt. The work of these authors was funded partially by the Hector Foundation II. We would also like to thank the team of "Wissenschaftliche Begleitung der Hector Kinderakademien" at the Hector Research Institute of Education Science and Psychology that generously provided us with constant research and administrative support.

References

Ambrósio, A. P., Xavier, C., & Georges, F. (2015). Digital ink for cognitive assessment of computational thinking. *Proceedings - Frontiers in Education Conference, FIE*, *February*. https://doi.org/10.1109/FIE.2014.7044237

Armoni, M. (2016). COMPUTING IN SCHOOLS Computer science, computational thinking, programming, coding. *ACM Inroads*, *7*(4), 24–27. https://doi.org/10.1145/3011071

Balanskat, A., & Engelhardt, K. (2015). *Computing our future*.

Barr, D., Harrison, J., & Conery, L. (2011). Computational Thinking: A Digital Age Skill for Everyone. *Learning and Leading with Technology*, *38*(6), 20–23.

Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12. *ACM Inroads*, *2*(1), 48–54. https://doi.org/10.1145/1929887.1929905

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(August), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

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

Bocconi, S., Chioccariello, A., Dettori, G., Ferrari, A., Engelhardt, K., Kampylis, P., & Punie, Y. (2016). Exploring the Field of Computational Thinking As a 21St Century Skill. *EDULEARN16 Proceedings*, *1*(June), 4725–4733. https://doi.org/10.21125/edulearn.2016.2136

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers and Education*, *94*(February), 178–192. https://doi.org/10.1016/j.compedu.2015.11.003

Brennan, K., & Resnick, M. (2012). Using artifact-based interviews to study the development of computational thinking in interactive media design. *Proceedings of the Annual American Educational Research Association Meeting (AERA)*. http://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf

Brooks, R. (1999). Towards a theory of the cognitive processes in computer programming. *International Journal of Human-Computer Studies*, *51*(2), 197–211. https://doi.org/doi.org/10.1006/ijhc.1977.0306

Butz, M. V. (2016). Toward a unified sub-symbolic computational theory of cognition. *Frontiers in Psychology*, *7*(JUN), 1–19. https://doi.org/10.3389/fpsyg.2016.00925

Butz, M. V., & Kutter, E. F. (2017). *How the Mind Comes into Being* (1st ed.). Oxford University Press.

Chen, G., Shen, J., Barth-Cohen, L., Jiang, S., Huang, X., & Eltoukhy, M. (2017). Assessing elementary students' computational thinking in everyday reasoning and robotics programming. *Computers & Education*, *109*, 162–175. https://doi.org/10.1016/j.compedu.2017.03.001

Citilab. (2015). *Scratch for Arduino*. http://s4a.cat/

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers.

Curzon, P., Dorling, M., Selby, C., Woollard, J., & Ng, T. (2014). *Developing computational thinking in the classroom: a framework*. *June*. http://eprints.soton.ac.uk/369594/10/DevelopingComputationalThinkingInTheClassroomaFramework.pdf

Dierbach, C., Hochheiser, H., Collins, S., Jerome, G., Ariza, C., Kelleher, T., Kleinsasser, W., Dehlinger, J., & Kaza, S. (2011). A Model for Piloting Pathways for Computational Thinking in a General Education Curriculum. *Development*, *15*(5), 257–262. https://doi.org/10.1145/1953163.1953243

Echeverría, A., García-Campo, C., Nussbaum, M., Gil, F., Villalta, M., Améstica, M., & Echeverría, S. (2011). A framework for the design and integration of collaborative classroom games. *Computers and Education*, *57*(1), 1127–1136. https://doi.org/10.1016/j.compedu.2010.12.010

Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS. (n.d.). *Open Roberta Lab*. https://lab.open-roberta.org/

García-Peñalvo, F. J., Reimann, D., Tuul, M., Rees, A., & Jormanainen, I. (2016). *TACCLE 3, O5: An overview of the most relevant literature on coding and computational thinking with emphasis on the relevant issues for teachers* (p. 54). Erasmus+.

Grover, S., & Pea, R. (2013). Computational Thinking in K-12: A Review of the State of the Field. *Educational Researcher*, *42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

Haffner, J., Baro, K., Parzer, P., & Resch, F. (2005). *HRT1-4: Heidelberger Rechentest; Erfassung mathematischer Basiskompetenzen im Grundschulalter.* Hogrefe.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? - A literature review of empirical studies on gamification. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 3025–3034. https://doi.org/10.1109/HICSS.2014.377

Hein, G. E. (1991). Constructivist Learning Theory. *CECA (International Committee of Museum Educators) Conference*. http://www.exploratorium.edu/ifi/resources/constructivistlearning.html

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision 3.* Beltz Test.

Herbein, E., Golle, J., Tibus, M., Zettler, I., & Trautwein, U. (2018). Putting a speech training program into practice: Its implementation and effects on elementary school children's public speaking skills and levels of speech anxiety. *Contemporary Educational Psychology*, *55*(September), 176–188. https://doi.org/10.1016/j.cedpsych.2018.09.003

Hickmott, D., & Prieto-Rodriguez, E. (2018). To assess or not to assess: Tensions negotiated in six years of teaching teachers about computational thinking. *Informatics in Education*, *17*(2), 229–244. https://doi.org/10.15388/infedu.2018.12

Hubwieser, P., & Mühling, A. (2014). *Play ying PISSA with Bebras*. 128–129.

Koh, K. H., Nickerson, H., Basawapatna, A., & Repenning, A. (2014). Early validation of computational thinking pattern analysis. *Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education - ITiCSE '14*, 213–218. https://doi.org/10.1145/2591708.2591724

Leifheit, L., Jabs, J., Ninaus, M., Moeller, K., & Ostermann, K. (2018). Programming unplugged: An evaluation of game-based methods for teaching computational thinking in primary school. *Proceedings of the European Conference on Games-Based Learning*, *2018-Octob*.

Liao, Y.-K. C., & Bright, G. W. (1991). Effects of Computer Programming on Cognitive Outcomes: A Meta-Analysis. *Journal of Educational Computing Research*, *7*(3), 251–268. https://doi.org/10.2190/E53G-HH8K-AJRR-K69M

Lifelong Kindergarten Group - MIT Media lab. (n.d.). *Scratch*. https://scratch.mit.edu/

Marinus, E., Powell, Z., Thornton, R., McArthur, G., & Crain, S. (2018). Unravelling the Cognition of Coding in 3-to-6-year Olds. *Proceedings of the 2018 ACM Conference on International Computing Education Research - ICER '18*, *August*, 133–141. https://doi.org/10.1145/3230977.3230984

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

Moreno-León, J., Robles, G., & Román-González, M. (2015). Dr. Scratch: Automatic Analysis of Scratch Projects to Assess and Foster Computational Thinking. *RED. Revista de Educación a Distancia*, *15*(46), 1–23. https://doi.org/10.6018/red/46/10

Mühling, A., Ruf, A., & Hubwieser, P. (2015). Design and First Results of a Psychometric Test for Measuring Basic Programming Abilities. *Proceedings of the Workshop in Primary and Secondary Computing Education*, 2–10. https://doi.org/10.1145/2818314.2818320

Ninaus, M., Pereira, G., Wood, G., Neuper, C., Stefitz, R., Prada, R., & Paiva, A. (2015). Game elements improve performance in a working memory training task. *International Journal of Serious Games*, *2*(1), 3–16. https://doi.org/10.17083/ijsg.v2i1.60

Noice, H., & Noice, T. (2001). Learning dialogue with and without movement. *Memory and Cognition*, *29*(6), 820–827. https://doi.org/10.3758/BF03196411

Parkinson, J., & Cutts, Q. (2018). *Investigating the Relationship Between Spatial Skills and Computer Science*. 106–114. https://doi.org/10.1145/3230977.3230990

Perkovic, L., Settle, A., Hwang, S., & Jones, J. (2010). A Framework for Computational Thinking across the Curriculum. *Proceedings of the Fifteenth Annual Conference on Innovation and Technology in Computer Science Education (ITiCS '10)*, 123–127. https://doi.org/10.1145/1822090.1822126

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of Game-Based Learning. *Educational Psychologist*, *50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Prottsman, K. (2014). Computer science for the elementary classroom. *ACM Inroads*, *5*(4), 60–63. https://doi.org/10.1145/2684721.2684735

Qualls, J. A., & Sherrell, L. B. (2010). Why computational thinking should be integrated into the curriculum. *Journal of Computing Sciences in Colleges*, *25*(5), 66–71.

Resnick, M. (2017). *Lifelong Kindergarten*. The MIT Press.

Román-González, M. (2015). Computational Thinking Test: Design Guidelines and Content Validation. *Proceedings of EDULEARN15 Conference*, *July 2015*, 2436–2444. https://doi.org/10.13140/RG.2.1.4203.4329

Román-González, M., Moreno-León, J., & Robles, G. (2017). Complementary Tools for Computational Thinking Assessment. *International Conference on Computational Thinking Education 2017*, *July*.

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

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

Sanders, M. (2009). STEM, STEM education, STEMmania. *The Technology Teacher*, *68*(4), 20–26.

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

Seiter, L., & Foreman, B. (2013). Modeling the learning progressions of computational thinking of primary grade students. *Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research - ICER '13*, 59. https://doi.org/10.1145/2493394.2493403

Settle, A., Franke, B., Hansen, R., Spaltro, F., Jurisson, C., Rennert-May, C., & Wildeman, B. (2012). *Infusing computational thinking into the middle- and high-school curriculum*. 22. https://doi.org/10.1145/2325296.2325306

Settle, A., Goldberg, D. S., & Barr, V. (2013). *Beyond computer science*. *July*, 311. https://doi.org/10.1145/2462476.2462511

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

Simon, H. A., & Newell, A. (2006). Human problem solving: The state of the theory in 1970. *American Psychologist*, *26*(2), 145–159. https://doi.org/10.1037/h0030806

Tsarava, K., Moeller, K., & Ninaus, M. (2018). Training Computational Thinking through board games: The case of Crabs & Turtles. *International Journal of Serious Games*, *5*(2), 25–44. https://doi.org/10.17083/ijsg.v5i2.248

Tsarava, K., Moeller, K., & Ninaus, M. (2019). Board Games for Training Computational Thinking. In M. Gentile, M. Allegra, & H. Söbke (Eds.), *Games and Learning Alliance* (pp. 90–100). Springer International Publishing.

Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2013). Challenges to learning and schooling in the digital networked world of the 21st century. *Journal of Computer Assisted Learning*, *29*(5), 403–413. https://doi.org/10.1111/jcal.12029

Weintrop, D., Beheshti, E., Horn, M. S., Orton, K., Trouille, L., Jona, K., & Wilensky, U. (2014). Interactive Assessment Tools for Computational Thinking in High School STEM Classrooms. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, *136 LNICST*, 22–25. https://doi.org/10.1007/978-3-319-08189-2_3

Weintrop, D., & Wilensky, U. (2015). Using commutative assessments to compare conceptual understanding in blocks-based and text-based programs. *ICER 2015 - Proceedings of the 2015*

ACM Conference on International Computing Education Research, *August*, 101–110. https://doi.org/10.1145/2787622.2787721

Weiß, R. H. (2006). *CFT 20-R: grundintelligenztest skala 2-revision.* Hogrefe.

Werner, L., Denner, J., & Campe, S. (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. https://doi.org/10.1145/2157136.2157200

Wiebe, E., Mott, B. W., London, J., Boyer, K. E., Aksit, O., & Lester, J. C. (2019). Development of a lean computational thinking abilities assessment for middle grades students. *SIGCSE 2019 - Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 456–461. https://doi.org/10.1145/3287324.3287390

Williamson, B. (2015). Political computational thinking: policy networks, digital governance and "learning to code." *Critical Policy Studies*, *10*(June), 1–20. https://doi.org/10.1080/19460171.2015.1052003

Wing, J. M. (2006). Computational Thinking. *Theoretical Computer Science*, *49*(3), 33–35. https://doi.org/https://www.cs.cmu.edu/~15110-s13/Wing06-ct.pdf

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. https://doi.org/10.1037/a0031311

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational Thinking in Elementary and Secondary Teacher Education. *ACM Transactions on Computing Education*, *14*(1), 1–16. https://doi.org/10.1145/2576872

Yadav, A., Zhou, N., Mayfield, C., Hambrusch, S., & Korb, J. T. (2011). Introducing Computational Thinking in Education Courses Aman. *Educational Studies*, *2*, 465–470. https://doi.org/10.1145/1953163.1953297

A Cognitive Approach to Defining and Assessing Computational Thinking: An Empirical Study in Primary School

Katerina Tsarava, Korbinian Moeller, Marcos Román-González, Jessika Golle, Luzia Leifheit, Martin V. Butz, Manuel Ninaus

Abstract

There is increasing effort on integrating Computational Thinking (CT) curricula, particularly but not limited to Computer Science Education (CSE). Therefore, research on the assessment of CT progressed towards the development and validation of reliable CT assessment tools necessary to evaluate students' learning gains due to developed curricular programs. Over the last years, several CT assessment tools were developed for elementary, high-school, and university students by which associations between CT scores and other cognitive abilities were unravelled. Like the general concept of *intelligence*, CT is only defined broadly as the ability to flexibly combine algorithmic operations to form complex solutions. In this study, we aimed at specifying a cognitive definition of CT, focusing on primary school level in a sample of 192 3rd and 4th graders. We used an adaptation of a validated CT test, which was initially designed for middle school students. Analyses indicated promising results regarding the reliability of the adapted CT assessment. Moreover, results revealed positive associations of CT with verbal reasoning-, non-verbal visuospatial-, and complex numerical abilities reflecting similarities but also differences between CT and other cognitive abilities, which essentially implies several basic cognitive abilities that support CT development differentially across time.

Keywords: 21st-century abilities, elementary education

1 Introduction

Computational thinking (CT) as a general problem-solving skill, along with others like communication, digital literacy, critical thinking, and creativity, has been coined an essential element of the so-called 21st-century skills (D. Barr et al., 2011; Bocconi et al., 2016; Voogt et al., 2013). The relevance of CT seems considerable, given the highly computerized world we live in. High demands for a digitally-trained workforce that develops and advances technology, on the one hand, and the need for well-informed and critical users of technology on the other, recently resulted in increasing educational and research interest in the early preparation of young students with and for digital technologies.

As Wing (2006) emphasized, when she introduced the term and concept, CT is a skill essential for everyone and not only for programmers or computer scientists. While computer programming draws on cognitive skills associated with CT and demonstrates computational competencies, CT can be applied to different kinds of problems that do not necessarily include programming (Román-González et al., 2018a). Even though the concept of CT is around for some time now, it still lacks a specific but widely accepted definition (though references to the term and attempts for elaborating the concept have been numerous, e.g., Curzon et al., 2014; Grover & Pea, 2013; National Research Council, 2011; Selby, 2013; Shute et al., 2017). As efforts for integrating CT in formal education increased over the last decade, the need for a better understanding and definition of CT becomes vital for the meaningful evaluation of new CT curricula, as well as for the development of relevant, accurate, and reliable CT assessment tools.

In the present study, we considered a recent definition of CT by Shute et al. (2017). In this literature review, Shute et al. specify CT as *“the conceptual foundation required to solve problems effectively and efficiently (i.e., algorithmically, with or without the assistance of computers) with solutions that are reusable in different contexts”* (Shute et al., 2017, p.142). Importantly, this definition portrays CT as a broader cognitive construct and not just a practical skill that is relevant for specific computing-related contexts (Armoni, 2016). It essentially emphasizes the broad applicability of CT as a universal problem-solving approach (Moreno-Leon et al., 2018).

Although the relevance of CT for computer science (CS) seems evident, CT as a cognitive construct has been considered beyond the field of CS (Armoni, 2016; Settle et al., 2013). The notion of CT as a 21st-century skill to be trained early on in education has led to the integration of CT in national curricula, either as a standalone learning field (Brown et al., 2014; Chiprianov & Gallon, 2016) or as an integral interdisciplinary part of other STEM topics (Dierbach et al., 2011; Perković et al., 2010; Settle et al., 2012; Weintrop, Beheshti, et al., 2016). Those efforts are continuously increasing (for a review, see Tang et al., 2020) and apply to a wide range of educational levels, from preschool (e.g., Bers et al., 2014; Sullivan et al., 2013) to primary (e.g., Duncan & Bell, 2015) and secondary school (e.g., Settle et al., 2012), to the university level (e.g., Dierbach et al., 2011).

1.1 Cognition of CT

In a recent meta-analysis, Scherer et al. (2019) provided valuable summative results from multiple empirical studies on the effects of programming – as a way of teaching, learning, and assessing computational thinking – on human cognition. These results showed near and far transfer effects on various cognitive skills (especially on creative thinking, mathematical skills, and reasoning). Although existing studies that are evaluating potential transfer effects of programming are mostly found on the primary school level (Scherer et al., 2019), studies on the cognitive correlates of programming are more focused on university or high school students (Román-González et al., 2018a). Even though considerable effort was devoted to studying CT as a cognitive construct [i.e., 80% of studies on CT assessment (Tang et al., 2020)], a comprehensive nomological network of cognitive abilities describing CT is still missing. In this study, we first review the available evidence for associations of CT with other cognitive abilities in primary school children. Notably, we consider associations with numerical/mathematical abilities, language, visuospatial abilities, and general cognitive abilities. Building upon the insights from our review, we designed our study to evaluate the cognitive correlates of CT, aiming at the development of a cognitive concept of CT for the underinvestigated age group of primary school students.

1.1.1 Computational thinking and numerical/mathematical cognition

A significant association of programming and numerical/mathematical cognition was first reported already a few decades ago (Pea & Kurland, 1984). This has been substantiated by further empirical evidence resulting from a plethora of studies since then (e.g., Byrne & Lyons, 2009; McCoy & Burton, 1988; Nowaczyk, 1983). For instance, mathematical cognition was observed to be a reliable predictor of performance in introductory programming courses (Bergin & Reilly, 2006). CT is an overarching cognitive construct that is essential for programming. Therefore, an evaluation of similarities but also differences between CT and numerical/mathematical cognition seems meaningful – in particular, because the latter refers to the application of numerical/mathematical abilities to solve mathematical but also more general problems in different STEM domains. In their work, Sneider et al. (2014) visualized the relationship between CT and numerical/mathematical thinking using a Venn diagram to indicate which capabilities might be considered part of numerical/mathematical cognition, CT, or both. In the same vein, Weintrop et al. (2016) concluded that the similarities between numerical/mathematical cognition and CT could be used for teaching mathematics and science classes in schools by using CT methods to support domain-specific learning.

However, results on the relationship between CT and numerical/mathematical cognition were inconsistent across different studies. For instance, Román-González et al. (2017) found no significant association between CT and numerical/mathematical abilities in a population of 1.251 10 to 16-year-old students. Similarly, a study with only a very small sample of 11 university students (Ambrosio et al., 2014) showed that calculus tasks did not significantly correlate with academic achievement in computer science. Nevertheless, in another study by Román-González et al. (2018) involving 314 12 to 13-year-old students, mathematics achievement correlated significantly but weakly with CT performance. Likewise, in a population of 348 university students, Werner (2019) observed a significant, but again only a weak correlation between CT and performance on simple arithmetic tasks like subtraction. However, they also found a significant weak correlation between CT and performance on more complex numerical tasks like algebra. In contrast, a previous study with a sample of 31 7 to 10-year-old primary school students showed a moderate to high correlation between CT and performance in simple and complex arithmetic tasks (Tsarava, Leifheit, et al., 2019). Similarly, in a recent study by Prat et al. (2020), involving 36 adults, numerical/mathematical abilities were a significant predictor of programming learning outcomes (i.e., learning rate, programming accuracy, and declarative knowledge).

In summary, previous evidence on the association of programming skills and CT with mathematics achievement and numerical/mathematical cognition provided inconsistent results across age. Interestingly, the pattern of results so far indicates a more pronounced association for younger children, which seems to decrease with age. Thus, basic numerical/mathematical abilities seem more relevant for CT in younger children attending primary school as compared to older ones, for whom more complex numerical/mathematical abilities seem to be more relevant. This could be explained by the fact that fundamental numerical/mathematical abilities, which are assessed by tests for younger children, are needed as prerequisites for thinking computationally. After a specific threshold of numerical/mathematical abilities is exceeded through formal education and due to the fact that most tests for older children and adolescents focus on curricular competencies rather than on fundamental numerical/mathematical abilities, the reviewed results indicate that these competencies may not be equally determining higher levels of CT (Tsarava, Leifheit, et al., 2019).

1.1.2 Computational thinking and language ability

Besides relationships to numerical and mathematical abilities, CT also revealed a partial relationship with language abilities. For instance, Marinus et al. (2018) found a positive association between cognitive compiling of syntax in natural language and programming ability in 28 3 to 6-year-old children. On the other hand, in an older population of 31 7 to 10-year-old children, Tsarava et al. (2019) did not observe a significant association between CT and performance on verbal reasoning. Howland & Good (2015) observed in a sample of 53 12 to 13-year-old children that the development of CT through programming activities is enhanced by activities that required the solution of narrative problems (e.g., digital storytelling) and a programming language that triggers verbal abilities by switching between formal and natural language, as a scaffolding strategy. Moreover, in a sample of 1.251 10 to 16-year-old students, a positive weak to moderate correlation between CT and verbal abilities was found (Román-González, Pérez-González, et al., 2017). In line with these results, Prat et al. (2020) found that for a sample of 36 adults between 18 and 35 years of age, language ability was a robust predictor of programming learning outcomes (i.e., learning rate, programming accuracy, and declarative knowledge) in a Python training course.

In summary, the majority of results indicate that language seems relevant for programming and CT. In one study (Román-González, Pérez-González, et al., 2017), even more than numerical/mathematical abilities.

1.1.3 Computational thinking and visuospatial abilities

For the last few years, empirical evidence has also accumulated for a mutual relation between CT and visuospatial abilities. For instance, in a primary school sample of 92 students (i.e., 1st to 5th graders), Città et al. (2019) found a stable positive association between visuospatial abilities (i.e., mental rotation) and CT as assessed by performance on a coding test. Likewise, results for a sample of 31 7 to 10-year-old primary school students showed that performance on visuospatial tasks significantly predicted CT (Tsarava, Leifheit, et al., 2019). Similar results were observed in a population of 1.251 10 to 16-year-old secondary school students by Román-González et al. (2017), who reported a moderate positive association between CT and spatial ability. Moreover, in an older population of 11 first-year university students, Ambrosio et al. (2014) reported a moderate to high association between academic achievement in programming and visuospatial reasoning. Substantiating these results, Werner (2020) also found a significant positive correlation between CT and performance on visuospatial tasks in a sample of 348 university students. Additionally, a study with 49 master's students (Jones & Burnett, 2008) again showed a positive association between mental rotation ability and attainment in a programming module. Finally, in a sample of students and academic staff of different levels of education in a CS department (n = 72), Parkinson & Cutts (2018) reported that the level of academic CS achievement increased with the visuospatial skills of participants.

These studies seem to indicate that visuospatial abilities assessed with either specific (e.g., mental rotation) or more general visuospatial tasks are associated significantly with CT – consistently from early primary school up to university level.

1.1.4 Computational thinking and general cognitive ability

CT, in the context of programming, is considered a problem-solving skill (e.g., Kalelioğlu et al., 2016). Generally, problem-solving skills seem closely related to aspects of fluid intelligence (Carroll, 1993), linked to non-verbal intelligence, as it is often assessed with language-free testing materials (Kubinger, 2009). In this context, a study with 28 3 to 6-year-old children (Marinus et al., 2018) found a positive association between programming ability and non-verbal intelligence. Furthermore, in a small-scale exploratory study with 12 first-year university students, Ambrosio et al. (2014) found a positive association between academic achievement in programming and general intelligence. Moreover, in a meta-analysis of 65 studies, Liao and Bright (1991) confirmed the association between general intelligence and learning a programming language. According to their findings, learning a programming language has positive effects on logical thinking and reasoning, as well as on the ability to plan and solve problems, regardless of which programming language was learned.

In contrast to programming skills, the cognitive correlates of specific CT skills have been explicitly investigated in a few studies only (e.g., Boom et al., 2018). For instance, a study on 1.251 10 to 16-year-old students (Román-González, Pérez-González, et al., 2017) found a positive small to moderate relationship between CT and three of the four primary cognitive abilities suggested by Thurstone (1939). Furthermore, they reported a high correlation between CT and children's problem-solving ability as a proxy of general cognitive ability (particularly fluid intelligence).

Taken together, these results provide strong empirical evidence for an association between CT and general cognitive abilities.

1.2 Assessment of CT

Assessment of CT has been pursued in various ways over the last few years (for a review on the latest research, see Román-González, Moreno-León, et al., 2017; Tang et al., 2020). There are assessment tools designed for specific educational programming environments, like Scratch (Brennan & Resnick, 2012b; Moreno-León et al., 2015; Seiter & Foreman, 2013), Alice (L. Werner et al., 2015; L. Werner, Denner, & Campe, 2012), and others (Koh, Basawapatna, et al., 2014), which primarily assess CT by evaluating projects created in the respective environment. Similarly, Weintrop et al. (2014) designed interactive digital assessments linked to specific STEM topics. Beyond these formative approaches to assess CT, there have been other attempts to develop psychometric tests assessing CT independent of specific programming environments.

An initial effort to assess CT psychometrically was taken by Mühling et al. (2015), who designed a summative assessment tool focusing on a specific programming concept. Alike, Weintrop and Wilensky (2015) designed a tool for assessing 5 different CT concepts (i.e., i. fundamentals like variables, assignment, etc., ii. selection statements, iii. definite loops, iv. indefinite loops, and v. function/method parameters). Similarly, Ambrósio et al. (2015) designed a digital cognitive assessment system aiming to assess CT, and particularly spatial reasoning, induction, and working memory. Another psychometric tool for assessing 5 circumscribed CT concepts (i.e., i. syntax, ii. data, iii. algorithms, iv. representations, and v. efficiency) was developed by Chen et al. (2017). One of the first attempts to assess CT in preschoolers was a test suggested by Marinus et al. (2018), which utilized a wooden programmable robot for implementing the assessment tasks. A common characteristic of all these propositions for the psychometric assessment of CT is their limited validation in larger samples and across different age levels.

To the best of our knowledge, the Computational Thinking test (*CTt*; Román-González, Pérez-González, et al., 2017) presents an exception to this as it has been validated systematically in the last few years (Román-González, 2015a; Román-González et al., 2018a, 2018b; Román-González, Moreno-León, et al., 2017; Wiebe et al., 2019). The *CTt* comprises 28 multiple-choice tasks designed to address a total of 7 different CT concepts (i. basic directions - sequences, ii. repeat times - loops, iii. repeat until - loops, iv. if - simple conditionals, v. if/else - complex conditionals, vi. while conditionals, and vii. simple functions) and was developed for the age group of 10 to 16-year-old children. At the time this study took place, no comparable assessment tool was available for our target population of 8 to 10-year-olds[11]. For that reason, we decided to adapt this assessment tool to create an abbreviated German version of the *CTt*, used to measure CT in 3rd and 4th graders.

1.3 Aim of the study

This study aimed at specifying the cognitive correlates of CT by evaluating its associations with specific cognitive abilities in primary school children. It builds upon the results of a pilot study and complements previous research conducted on other age groups, such as preschoolers (e.g., Marinus et al., 2018), middle school (e.g., Román-González, Pérez-González, & Jiménez-Fernández, 2017), and university students (e.g., Ambrósio et al., 2011). Additionally, the present study presents a new tool for assessing CT as a unique cognitive ability in primary school students, by adapting an assessment test initially designed for older students (*CTt*; Román-González, Pérez-González, & Jiménez-Fernández, 2017).

[11] A new assessment similar to *CTt* but not yet developed at the time this study took place is the Computational Thinking Test for Beginners (*BCTt*; Zapata-Cáceres et al., 2020), which follows a similar concept design approach as the *CTt*, but uses simplified and friendlier to young children task elements. Similarly, Relkin et al. (2020) recently proposed an unplugged CT assessment for 5 to 9-year-old students in a multiple-choice format.

2 Method

2.1 Participants

This study was conducted within the Hector Children's Academy Program (HCAP) in the German state of Baden- Württemberg. HCAP provides extracurricular enrichment courses at 66 local sites that offer additional courses for primary school students (grades 1 through 4; for details on the recruitment process and HCAP, see Rothenbusch et al. 2016).

Overall, students enrolled in 16 different local sites of the HCAP participated voluntarily. Each local site had 6 to 20 participating students in 1 or 2 groups of 6 to 10 students each. Overall, 196 third- (n = 85; 72,9% male; age: M = 8.64, SD = 0.48 years) and fourth-grade (n = 111; 78.40% male; age: M = 9.62, SD = 0.44 years) elementary school students participated in the study. Data of 192 children were submitted to further analyses (75.50% male; age: M = 9.20, SD = 0.67 years). Prior to the study, we obtained written, informed consent from the legal guardians of the students and the students themselves. The study was approved by the Ethics Committee of the Leibniz Institut für Wissensmedien.

2.2 Measures

To evaluate the association between CT and relevant other cognitive abilities, we administered various scales of standardized test batteries. In addition to these measurements, we used two self-developed scales for the assessment of short-term visual and auditory/verbal memory. Preliminary analysis showed insufficient psychometric properties for these two scales in comparison to the established standardized tests. Therefore, these scales were not considered in the present analyses.

In a 120-minutes long session, we assessed students on the following pen-and-paper cognitive tests:

2.2.1 Numerical /mathematical abilities

For the assessment of numerical/mathematical abilities, we administered four speeded subtests of the *Heidelberger Rechentest zur Erfassung mathematischer Basiskompetenzen im Grundschulalter* (HRT 1-4; Haffner, 2005): i) *Addition* (2 minutes), ii) *Subtraction* (2 minutes), iii) *Multiplication* (fact retrieval; 2 minutes), and iv) *Problem Completion* (2 minutes). Each of the scales, *Addition, Subtraction,* and *Multiplication,* comprises 40 problems, and participants are asked to correctly solve as many as possible within the time limit of 2 minutes. Each of the scales begins with simple one-digit problems (e.g. 2+1, 3-1, 3*2) that gradually increase in difficulty through involving more digits (e.g. 366 + 512, 531 - 274, 12 * 15). The scale *Problem Completion* contains 40 mixed addition and subtraction problems, where one operand is missing (Haffner et al., 2005). Again, the scale begins with simple problems (e.g., 14 = 8 + _) gradually increasing in difficulty. (e.g. 45 -12 = 21 + _).

Additionally, the subscale *Writing Speed* of the HRT was employed as a measure of processing speed. When completing the *Writing Speed* scale, participants are asked to copy as many target numbers as possible from a total of 40 1-digit numbers within the time limit of 30 seconds.

2.2.2 Non-verbal visuospatial reasoning

Non-verbal visuospatial reasoning was assessed using two speeded subtests of the *Culture Fair Intelligence Test* (CFT 20-R; Weiß, 2006) that comprises inductive reasoning tasks, namely the *Series Continuation* (Subtest 1) and *Matrices* (Subtest 3). CFT utilizes language-independent graphic tasks to measure the ability to identify patterns and relationships. Participants had 7 minutes to complete both subtests that contain 15 items each. On each item of the *Series Continuation* scale, a series of three figures is presented. From a set of five alternatives, participants have to choose a fourth matching figure that completes the progression of the series. On each item of the *Matrices* scale, an incomplete 2x2 matrix of figures is presented. From a set of five alternatives, participants have to choose the figure that complies with the respective underlying rules.

2.2.3 Verbal reasoning

Verbal reasoning was assessed using the speeded subtest *Vocabulary* (V Test 1-Form A) of the *Kognitiver Fähigkeitstest* (KFT 4-12+R; Heller & Perleth, 2000), which is the German version of the Cognitive Abilities Tests (CAT; Thorndike & Hagen, 1971). For completing the 25 items of this scale, participants had 7 minutes. For this scale, students are presented with a word and are asked to choose another word from five given alternatives that has the same or at least the most similar meaning (e.g., a synonym or a generic umbrella term).

2.2.4 Computational thinking

CT was assessed using 21 items of the *Computational Thinking test* (CTt) (Román-González, Pérez-González, et al., 2017), which were translated into German. We used an abbreviated version of the CTt scale. In particular, we used the 21 easiest items out of the total number of 28 items that the original test comprises (i.e., original item numbers: 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 13, 14, 17, 18, 19, 20, 21, 24, 26, 27, 28). These items were selected because the sample assessed in the current study was much younger as compared to the original sample for which the test was validated. The *abbreviated CTt* can be found in German on the Open Science Framework (OSF) platform[12]. The seven different CT concepts addressed by the test are i. *basic directions* and *sequences*, ii. *loops* implemented with *repeat-times* commands, iii. *loops* implemented with *repeat-until* commands, iv. *simple conditionals* implemented with *if* commands, v. *complex conditionals*, implemented with *if/else* commands, vi. w*hile conditionals*, and vii. simple *functions*. Test items are presented in the form of a 'Pac-Man' maze or an artist's canvas and request responses in the form of visual arrows or visual programming blocks. Participants are required to navigate an object from one place to another on a 2D environment using different commands in the form of programming blocks or directive arrows.

3 Results

3.1 Abbreviated CTt

The main descriptive statistics of the *abbreviated CTt* score for our sample are shown in *Table 1*. Generally, they are comparable to the descriptive statistics of the original *CTt*, as described by Román-González et al. (2017). All analyses were performed with IBM SPSS (version 25) and R (version 4.0.0). The distribution of the *abbreviated CTt* along the sample is depicted in Figure 1. and seems to approach normal distribution. To evaluate sex differences in performance on the *CTt*, we ran an independent t-test (t (190) = 1.790, p = .075, d =.311). These results indicate no significant sex difference in *CTt* performance. Although the difference was not statistically significant, the mean *CTt* performance of girls (m = 10.64, SD = 3.05) was lower than that of boys (m = 11.66, SD = 3.48).

[12] The temporary web link to the OSF repository is https://osf.io/4tp9c/?view_only=23f90904989c471991f429fd56972c61

Table 1. Descriptive statistics of the abbreviated CTt.

	N	Mean	Std. Error of Mean	Std. Deviation	Variance	Skewness	Kurtosis	Min	Max
CTt	192	11.41	.246	3.403	11.583	-.050	-.580	3	19

The reliability of the *abbreviated CTt*, as indicated by Cronbach's alpha, was *α* = .637. To overcome concerns about the alpha coefficient and its assumptions as a biased measurement of internal consistency, we decided to also compute the omega alongside a confidence interval (Dunn et al., 2013). This allows for a more accurate consistency estimation in case the *abbreviated CTt* is not assessing a unidimensional structure. The coefficient omega was .638, 95% CI [.55, .71, *SE* = 0.04]. Both measurements seem comparable and indicate adequate reliability of the scale, considering the young population it was administered to and its speeded administration.

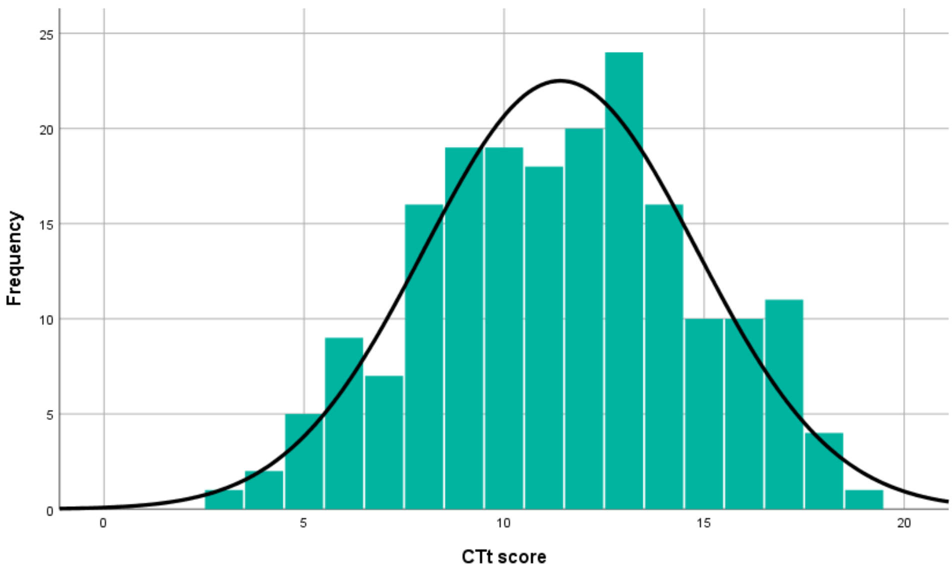

Figure 1. Histogram of participants' scores on the abbreviated CTt.

To evaluate whether the construct measured by the *abbreviated CTt* is unidimensional and, therefore, correctly interprets reliability as indexed by Cronbach's *α,* we performed an exploratory factor analysis. Inspection of the scree plot (see Figure 2) and based on the so-called elbow criterion, results of the exploratory factor analysis indicated a one-factor solution. As such, our results are in line with those reported by Román-González (2016) on the *CTt* with 28 items assessed in 10 to16-year-old children. Similar results on the unidimensionality of the *CTt* have also been reported by Guggemos et al. (2019) and Wiebe et al. (2019).

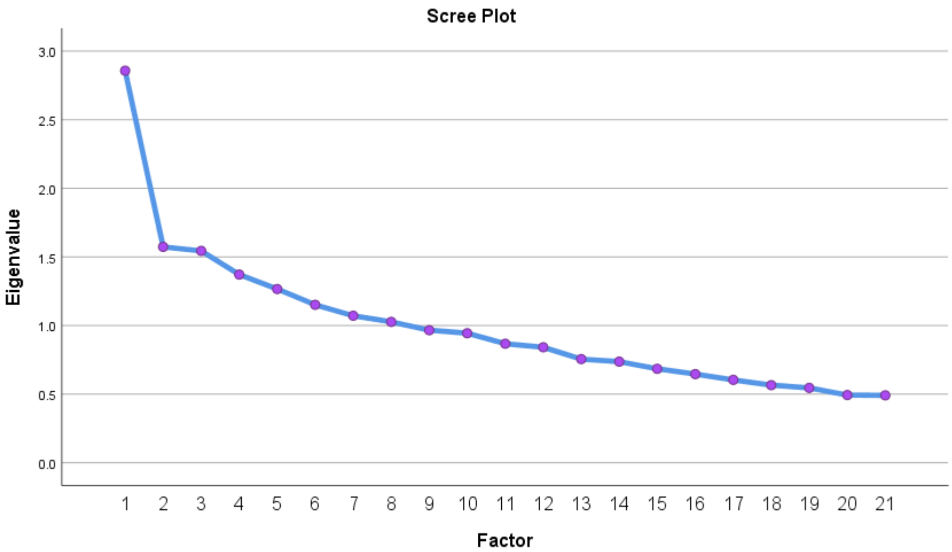

Figure 2. Scree plot of the factor analysis of the abbreviated CTt.

3.2 Correlations

To examine associations between CT and specific other cognitive abilities, the same analysis procedure as in (Tsarava, Leifheit, et al., 2019) was followed by first running correlation analyses between the following variables:

- WRS: number of correct responses in the *Writing Speed* subtest of the HRT 1-4. The maximum score value is 40.
- CALC: mean of correct responses in the *Addition* and *Subtraction* subtests of the HRT 1-4. The maximum score value for *Calculation* is 40.
- MUL: number of correct responses in the *Multiplication* subtest of the HRT 1-4. The maximum score value is 40.
- COM: number of correct responses in the *Problem Completion* subtest of the HRT 1-4. The maximum score value is 40.
- CFT: mean of correct responses in the *Series Continuation* and *Matrices* subtests of the CFT. The maximum score value is 30.
- KFT: number of correct responses in the *Vocabulary* subtest of the KFT. The maximum score value is 25.

- *CTt*: number of correct responses in the shortened version of *CTt*. The maximum score value is 21.

The *CTt* performance showed a weak positive correlation with grade (r =.324, n = 192, p < .001) and age (r = .286, n = 192, p < .001). These results suggest that CT abilities increase with the grade and the age of the students assessed.

Results of the correlation analyses between *CTt* performance and the other cognitive tests (WRS, CALC, MUL, COM, CFT, KFT) are reported in *Table 2*. There was a moderately positive correlation between performance at the *CTt* and the KFT *Vocabulary* subtest (r = .388, n = 191, p < .001). Additionally, we observed significant positive correlations between *CTt* performance and performance on the CFT (r = .346, n = 192, p < .001), the HRT *Problem Completion* subtest (r = .333, n = 192, p < .001), the HRT *Multiplication* subtest (r = .224, n = 192, p = .002), and CALC performance (Calculation abilities reflecting performance on HRT subscales *Addition* and *Subtraction*, r = .232, n = 192, p = .001).

Corresponding scatter plots are shown in Figure 3. There was no significant correlation between *CTt* performance and the HRT *Writing Speed* subtest. These results suggest that CT is primarily related to verbal and non-verbal visuospatial reasoning abilities, more complex arithmetic processes like problem completion, and secondarily with basic arithmetic operations like addition, subtraction, and multiplication.

Table 2. Correlations (Pearson's r) between CT performance and other cognitive abilities scores.

	WRS	CALC	MUL	COM	CFT	KFT
CTt	-.023	.232**	.224**	.333**	.346**	.388**
WRS		.297**	.328**	.244**	.098	.076
CALC			.714**	.621**	.308**	.307**
MUL				.546**	.273**	.187**
COM					.325**	.306**
CFT						.397**
KFT						

**. Correlation is significant at the 0.01 level (2-tailed).

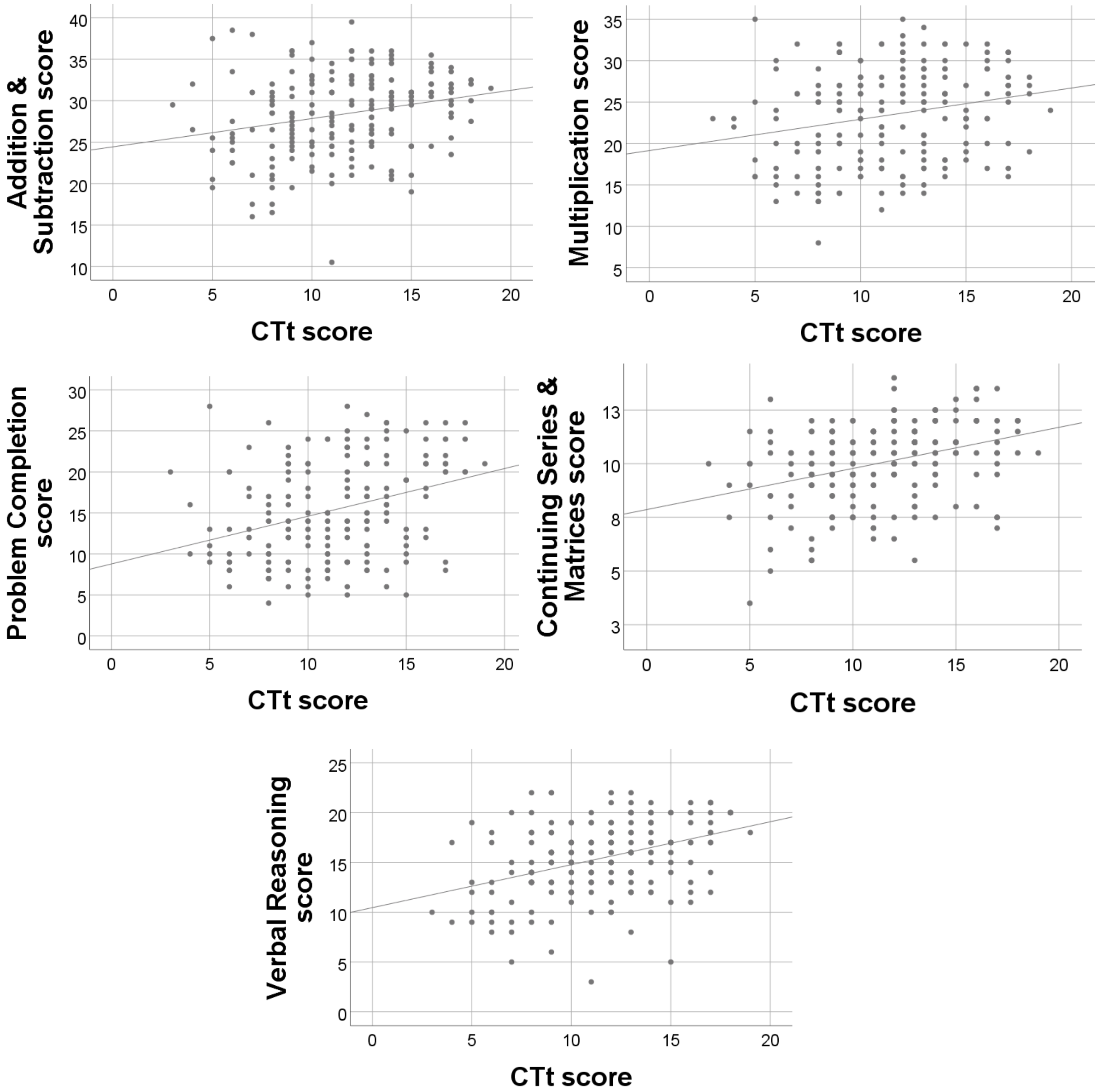

Figure 3. Scatter plots reflecting association of CTt performance score and other cognitive performance scores.

3.3 Regression

A multiple linear regression analysis utilizing the ENTER method was performed to investigate the prediction of *CTt* performance by specific cognitive skills. To do so, verbal reasoning, problem completion, visuospatial reasoning, writing speed, calculation abilities, and multiplication performance were used as predictors. The multiple regression model significantly predicted *CTt* performance and explained about 24% of variance (F (6, 179) = 9.635; p < .001, R^2 = .244, adj. R^2 = .219). In particular, *CTt* performance was predicted significantly by verbal reasoning, problem completion, and non-verbal visuospatial reasoning.

Inspection of beta weights indicated that better *CTt* performance was predicted by better verbal reasoning (KFT; β = .265, p < .001), better problem completion ability (COM; β = .214, p = .014), and better visuospatial reasoning (CFT; β = .178, p = .016). For *Writing speed*, a tendency was observed with better writing speed tending to predict better *CTt* performance (WRS; β = -.124, p = .074). Calculation abilities (CALC; β = -.073, p = .482) and multiplication (MUL; β = .101, p = .296) skills were no significant predictors of *CTt* performance.

4 Discussion

Over the last 15 years, numerous national and international initiatives suggested CT as a crucial 21st-century skill (for a review, see Hsu et al., 2019). Research so far supports the notion that CT seems to be a skill that should be taught early on in education by means of an interdisciplinary curriculum including programming activities, mathematics, and other STEM and non-STEM subject domains (Curzon et al., 2014; Perković et al., 2010; Shute et al., 2017). However, the field is still in its infancy. A clear and commonly accepted definition of CT grounded on empirical evidence is still missing. Yet, this definition is vital for the development and evaluation of effective instructional materials that help to foster CT as well as for a reliable assessment of CT skills and curricula.

To advance a cognitive definition of CT, the current study investigated the cognitive correlates of CT, focusing on analyzing association with established assessments of other cognitive abilities in primary school students – a sample that has been fairly neglected so far in CT research. We, therefore, modified an existing CT assessment tool for older students to make it appropriate for the younger target population. We selected a subset of items by omitting more difficult items to make it appropriate for the younger target population. Considering the young age group and its administration as a speeded test, the *abbreviated CTt* showed acceptable reliability both in a pilot (Tsarava, Leifheit, et al., 2019) and the current study. Importantly though, overall, our *abbreviated CTt* for primary school students showed similar psychometric properties as the original CT intended for older students. Additionally, we observed significant associations between *CTt* and numerical abilities, verbal abilities, and non-verbal visuospatial reasoning abilities. In fact, 24% of variance of *CTt* was explained by these cognitive abilities. Taken together, this means that *CTt* is associated only weakly to moderately with other cognitive abilities, which in turn provides further evidence for CT to be a specific cognitive ability that seems to build on and recruit a convolute of several other cognitive abilities (for a summative visualization of the correlational analysis, see Figure 4). Importantly, please note that the cognitive abilities for which we evaluated potential associations with CT in this study are not exhaustive, and there may be other abilities, like creativity, working memory, etc. that also convolute in the development of CT. In the following, we will discuss our results with respect to other findings in the field.

Positive association between CT and numerical/mathematical cognition: The weak positive association observed between CT and simple as well as complex numerical/mathematical abilities replicated findings of a previous study using the same materials (Tsarava, Leifheit, et al., 2019). Although studies on older middle school students did not provide similar evidence (Román-González, Pérez-González, et al., 2017), these results substantiate the argument that at an early stage of development, numerical abilities may be a prerequisite for thinking computationally. However, after a certain threshold of numerical abilities may be reached, for instance, by formal education, advanced numerical abilities (e.g., algebra, geometry, etc.) may not be mandatory to develop CT further. This claim is in line with recent empirical evidence indicating that also in adult populations; numeracy is not the most pronounced predictor of thinking computationally in the context of programming (Helmlinger et al., 2020; Prat et al., 2020).

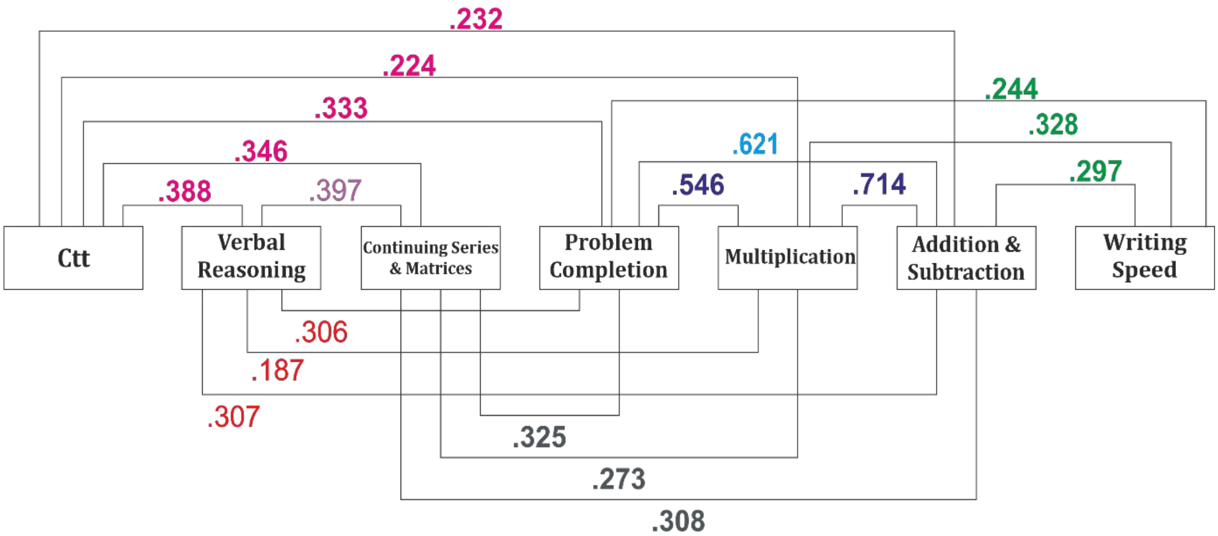

Figure 4. Tree structure of the significant correlational results between CTt performance score and all the other cognitive performance scores.

Positive association of CT and language abilities: The moderate positive association between CT and verbal reasoning that was observed in this study is well in line with the majority of previous studies conducted in older (e.g., Román-González, Pérez-González, & Jiménez-

Fernández, 2017) but also younger children (Marinus et al., 2018). Furthermore, the current findings substantiate results found in studies investigating the association of programming abilities and language abilities. In sum, these results provide convincing evidence on a significant association of CT with language abilities in primary school children (8 to 10 years old), similar to the ones already observed in preschool (3 to 6 years old; Marinus et al., 2018) and middle school students (10 to 16 years old; Román-González, Pérez-González, & Jiménez-Fernández, 2017) as well as in young adults (Prat et al., 2020). For the latter population, verbal ability is not associated with programming experience and CT, as shown by a recent study (Helmlinger et al., 2020). This decreasing association of CT with verbal ability with age and after secondary education might be explained by the use of language as a scaffolding cognitive strategy to read and write algorithms, which becomes less relevant with age. Nonetheless, as regards language, there is empirical evidence for the effectiveness of narrative and storytelling activities on promoting CT (e.g., Howland & Good, 2015). The *CTt* utilized in this study requires verbal instruction of the tasks and partially a verbally structured solution (e.g., with programming blocks). The involvement of such verbal demands might explain the association of CT performance and verbal ability found in this and previous studies.

Positive association of CT with visuospatial abilities: The weak positive association between CT and visuospatial abilities substantiates our previous findings (Tsarava, Leifheit, et al., 2019) as well as the findings of numerous other studies conducted on different age groups (like Città et al. (2019) in primary school students; (Román-González, Pérez-González, & Jiménez-Fernández (2017) in middle school; Jones & Burnett, (2008) in master's students; and others (Ambrosio et al., 2014; Parkinson & Cutts, 2018; M. Werner, 2020). In the present study, visuospatial processing was a relevant predictor of CT performance. This might reflect that the *CTt* is quite demanding with respect to visuospatial abilities (e.g., spatial navigation through mazes and on the grid). In sum, it seems that non-verbal visuospatial reasoning abilities are a consistent predictor of thinking computationally across educational levels, from early primary school to university level.

Contextualizing the results of this study within the literature on the existing studies on CT's cognitive correlates across educational levels, three main conclusions may be drawn from a developmental perspective (for a visualization, see Figure 5). First, basic numerical abilities (like addition, subtraction, and multiplication) seem to be prerequisites for developing CT at primary school level, whereas this seems not the case later on in education after a specific threshold of mathematical ability has been achieved. Second, verbal abilities seem to be relevant for developing CT both along primary and secondary education levels. It may well be that they are relevant for verbalizing an algorithmic process as a scaffolding cognitive strategy at an early stage. However, such semantic mapping of words to actions may be less important later on after secondary school. Third, non-verbal reasoning abilities seem to be essential for CT all the way from primary education level up to the university level and beyond. Importantly, non-verbal inductive reasoning often reflects some kind of pattern recognition and abstraction, two main elements of CT. As non-verbal inductive reasoning is usually considered the main element of fluid intelligence tests, CT might consequently also be considered a general problem-solving ability.

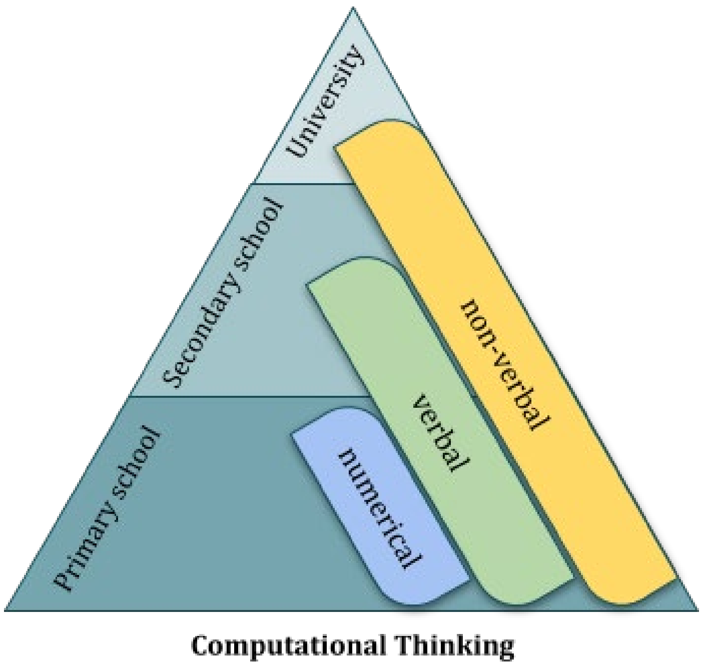

Figure 5. Development of CT across educational levels.

Taken together, results of the current correlation and regression analyses clearly suggest that CT in primary school children is primarily related to verbal reasoning abilities, visuospatial reasoning, and more complex arithmetic processes. Nevertheless, these cognitive correlates of CT explain about 24% of the variance in CT performance only. As such, these results indicate that CT is only weakly to moderately associated with other cognitive abilities. In turn, this corroborates the assumption of CT representing a special and specific cognitive ability. Future studies are needed to further substantiate the acclamation of CT as a unique cognitive ability that relies on a convolute of several other cognitive abilities, taking into consideration factors not considered in this study, like creative thinking (Scherer et al., 2018), executive functions (Robertson et al., 2020), and non-cognitive behavioural factors, like personality or self-efficacy (Román-González et al., 2018b).

Interestingly, in contrast to the initial validation of the *CTt* for secondary-school students (Román-González, Pérez-González, et al., 2017), we observed no significant sex differences in *CTt* performance. This result may support the argument of Roman-Gonzales et al. that sex differences in CT seem to increase as children grow older and advance to higher educational levels. This pattern has already been observed in the development of other cognitive skills as well (e.g., Keith et al., 2008). However, the result of no significant sex differences in *CTt* performance should be interpreted with caution, as our sample only comprised a smaller subsample of girls compared to boys. Comparing means of performance for the two subsamples showed better performance for boys than for girls, which is consistent with what Román-González, Pérez-González, & Jiménez-Fernández (2017) observed on an older sample.

4.1 Limitations

When interpreting the results of the current study, it needs to be acknowledged that the participating students had previously been nominated by their teachers to attend an extracurricular enrichment program based on their school achievement and/or motivation for the topic. Consequently, our sample may not be representative of the overall student population but represents a sample biased towards better-performing students. This should be kept in mind when generalizing our results beyond the current population.

5 Conclusions and Further Research

In this work, we largely replicated the so far observed positive associations of CT with other cognitive measures (i.e., language and visuospatial abilities) in a sample of 192 8 to 10-year-old primary school children. Additionally, we observed a positive association of CT with numerical/mathematical abilities, which was not observed in similar studies with secondary school children. Our regression results on the cognitive correlates of CT explained 24% of the variance in CT performance, which supports the assumption that CT seems to be a unique cognitive ability, which needs further empirical investigation. Additionally, we have adapted an existing CT assessment tool for secondary school children, making the *abbreviated CTt* appropriate for primary school children. Thus, we provide evidence on its reliability for measuring CT in the respective age group.

Future work shall focus on the evaluation of CT curricula utilizing the *abbreviated CTt* along with other cognitive tests to investigate the transfer effects of CT trainings. Moreover, further cross-validation studies of the *abbreviated CTt* in comparison to newly developed assessment tools like the *BCTt* (Zapata-Cáceres et al., 2020) and well established Bebras tasks (Dagiene & Stupuriene, 2016) are desirable to ensure its psychometric quality. Nevertheless, the current study demonstrated the applicability of the *abbreviated CTt* in young children of 8 to 10 years of age.

ACKNOWLEDGEMENTS

We gratefully acknowledge the help provided by our student research assistants Jana Hofmann, Moritz Werner, Joachim Fritscher, Joshua Schmid, and Daniela Piechnik. Thanks are also due to the bachelor's students Christian Reiff, Mareike Nutz, Marcel Fröhlich, and Ioana Uhl. Besides, we are grateful to all the teachers and headmasters of the 16 Hector Children's Academies that participated in the study, as well as the team "Wissenschaftliche Begleitung der Hector Kinderakademien" at the Hector Research Institute of Education Sciences and Psychology, that coordinated this study and provided us with constant research feedback and administrative assistance. This study was made possible by funding of the Hector Foundation II.

References

Ambrosio, A. P., Costa, F. M., Almeida, L., Franco, A., & Macedo, J. (2011). Identifying cognitive abilities to improve CS1 outcome. *2011 Frontiers in Education Conference (FIE)*, *February*, F3G-1-F3G-7. https://doi.org/10.1109/FIE.2011.6142824

Ambrosio, A. P., da Silva Almeida, L., Macedo, J., & Franco, A. (2014). Psychology of Programming Interest Group Annual Conference 2014. *PPIG Proceeding*, *July*, 1–25. http://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf%0Ahttp://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.698.1911&rep=rep1&type=pdf

Ambrósio, A. P., Xavier, C., & Georges, F. (2015). Digital ink for cognitive assessment of computational thinking. *Proceedings - Frontiers in Education Conference, FIE*, *February*. https://doi.org/10.1109/FIE.2014.7044237

Armoni, M. (2016). COMPUTING IN SCHOOLS Computer science, computational thinking, programming, coding. *ACM Inroads*, *7*(4), 24–27. https://doi.org/10.1145/3011071

Barr, D., Harrison, J., & Conery, L. (2011). Computational Thinking: A Digital Age Skill for Everyone. *Learning and Leading with Technology*, *38*(6), 20–23.

Bergin, S., & Reilly, R. (2006). Predicting introductory programming performance: A multi-institutional multivariate study. *Computer Science Education*, *16*(4), 303–323. https://doi.org/10.1080/08993400600997096

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

Bocconi, S., Chioccariello, A., Dettori, G., Ferrari, A., Engelhardt, K., Kampylis, P., & Punie, Y. (2016). Exploring the Field of Computational Thinking As a 21St Century Skill. *EDULEARN16 Proceedings*, *1*(June), 4725–4733. https://doi.org/10.21125/edulearn.2016.2136

Boom, K. D., Bower, M., Arguel, A., Siemon, J., & Scholkmann, A. (2018). Relationship between computational thinking and a measure of intelligence as a general problem-solving ability. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, 206–211. https://doi.org/10.1145/3197091.3197104

Brennan, K., & Resnick, M. (2012). Using artifact-based interviews to study the development of computational thinking in interactive media design. *Proceedings of the Annual American Educational Research Association Meeting (AERA)*. http://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf

Brown, N. C. C., Sentance, S. U. E., Crick, T. O. M., & Humphreys, S. (2014). Restart: The Resurgence of Computer Science in UK Schools. *ACM Transactions on Computing Education*, *14*(2), 1–22. https://doi.org/10.1145/2602484

Byrne, P., & Lyons, G. (2001). The effect of student attributes on success in programming. *ACM SIGCSE Bulletin*, *33*(3), 49–52. https://doi.org/10.1145/507758.377467

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. Cambridge University Press. https://doi.org/10.1177/001698629904300207

Chen, G., Shen, J., Barth-Cohen, L., Jiang, S., Huang, X., & Eltoukhy, M. (2017). Assessing elementary students' computational thinking in everyday reasoning and robotics programming. *Computers & Education*, *109*, 162–175. https://doi.org/10.1016/j.compedu.2017.03.001

Chiprianov, V. (2016). Introducing Computational Thinking to K-5 in a French Context. *ITiCSE*. https://doi.org/10.1145/2899415.2899439

Città, G., Gentile, M., Allegra, M., Arrigo, M., Conti, D., Ottaviano, S., Reale, F., & Sciortino, M. (2019). The effects of mental rotation on computational thinking. *Computers & Education*, *141*(July), 103613. https://doi.org/10.1016/j.compedu.2019.103613

Curzon, P., Dorling, M., Selby, C., Woollard, J., & Ng, T. (2014). *Developing computational thinking in the classroom: a framework*. *June*. http://eprints.soton.ac.uk/369594/10/DevelopingComputationalThinkingInTheClassroomaFramework.pdf

Dagiene, V., & Stupuriene, G. (2016). Bebras - A sustainable community building model for the concept based learning of informatics and computational thinking. *Informatics in Education*, *15*(3), 25–44. https://doi.org/10.15388/infedu.2016.02

Dierbach, C., Hochheiser, H., Collins, S., Jerome, G., Ariza, C., Kelleher, T., Kleinsasser, W., Dehlinger, J., & Kaza, S. (2011). A Model for Piloting Pathways for Computational Thinking in a General Education Curriculum. *Development*, *15*(5), 257–262. https://doi.org/10.1145/1953163.1953243

Duncan, C., & Bell, T. (2015). A pilot computer science and programming course for primary school students. *ACM International Conference Proceeding Series*, *09-11-Nove*, 39–48. https://doi.org/10.1145/2818314.2818328

Dunn, T. J., Baguley, T., & Brunsden, V. (2013). From Alpha to Omega. *The British Journal of Psychology*, *105*(3), 399–412.

Grover, S., & Pea, R. (2013). Computational Thinking in K-12: A Review of the State of the Field. *Educational Researcher*, *42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

Guggemos, J., Seufert, S., & Román-González, M. (2019). Measuring computational thinking - Adapting a performance test and a self-assessment instrument for German-speaking countries. *16th International Conference on Cognition and Exploratory Learning in Digital Age, CELDA 2019*, *Celda*, 183–191.

Haffner, J., Baro, K., Parzer, P., & Resch, F. (2005). *HRT1-4: Heidelberger Rechentest; Erfassung mathematischer Basiskompetenzen im Grundschulalter.* Hogrefe.

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision 3.* Beltz Test.

Helmlinger, B., Sommer, M., Feldhammer-Kahr, M., Wood, G., Arendasy, M. E., & Kober, S. E. (2020). Programming experience associated with neural efficiency during figural reasoning. *Scientific Reports*, *10*(1), 1–14. https://doi.org/10.1038/s41598-020-70360-z

Howland, K., & Good, J. (2015). Learning to communicate computationally with Flip: A bi-modal programming language for game creation. *Computers and Education*, *80*, 224–240. https://doi.org/10.1016/j.compedu.2014.08.014

Hsu, Y. C., Irie, N. R., & Ching, Y. H. (2019). Computational Thinking Educational Policy Initiatives (CTEPI) Across the Globe. *TechTrends*, 260–270. https://doi.org/10.1007/s11528-019-00384-4

Jones, S., & Burnett, G. (2008). Spatial Ability and Learning to Program. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, *4*(1), 47–61. https://doi.org/10.17011/ht/urn.200804151352

Kalelioğlu, F., Gülbahar, Y., & Kukul, V. (2016). A Framework for Computational Thinking Based on a Systematic Research Review. *Baltic J . Modern Computing*, *4*(3), 583–596.

Keith, T. Z., Reynolds, M. R., Patel, P. G., & Ridley, K. P. (2008). Sex differences in latent cognitive abilities ages 6 to 59: Evidence from the Woodcock-Johnson III tests of cognitive abilities. *Intelligence*, *36*(6), 502–525. https://doi.org/10.1016/j.intell.2007.11.001

Koh, K. H., Basawapatna, A., Nickerson, H., & Repenning, A. (2014). Real time assessment of computational thinking. *Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC*, 49–52. https://doi.org/10.1109/VLHCC.2014.6883021

Marinus, E., Powell, Z., Thornton, R., McArthur, G., & Crain, S. (2018). Unravelling the Cognition of Coding in 3-to-6-year Olds. *Proceedings of the 2018 ACM Conference on International Computing Education Research - ICER '18*, *August*, 133–141. https://doi.org/10.1145/3230977.3230984

McCoy, L. P., & Burton, J. K. (1988). The relationship of computer programming and mathematics in secondary students. *Computers in the Schools*, *4*(3–4), 159–166.

Moreno-León, J., Robles, G., & Román-González, M. (2015). Dr. Scratch: Automatic Analysis of Scratch Projects to Assess and Foster Computational Thinking. *RED. Revista de Educación a Distancia*, *15*(46), 1–23. https://doi.org/10.6018/red/46/10

Moreno-Leon, J., Roman-Gonzalez, M., & Robles, G. (2018). On computational thinking as a universal skill: A review of the latest research on this ability. *2018 IEEE Global Engineering Education Conference (EDUCON)*, 1684–1689. https://doi.org/10.1109/EDUCON.2018.8363437

Mühling, A., Ruf, A., & Hubwieser, P. (2015). Design and First Results of a Psychometric Test for Measuring Basic Programming Abilities. *Proceedings of the Workshop in Primary and Secondary Computing Education*, 2–10. https://doi.org/10.1145/2818314.2818320

National Research Council. (2011). *Report of a Workshop of Pedagogical Aspects of Computational Thinking*. https://doi.org/978-0-309-21474-2

Nowaczyk, R. H. (1983). *Cognitive Skills Needed in Computer Programming*. https://www.learntechlib.org/p/136288

Parkinson, J., & Cutts, Q. (2018). *Investigating the Relationship Between Spatial Skills and Computer Science*. 106–114. https://doi.org/10.1145/3230977.3230990

Pea, R. D., & Kurland, D. M. (1984). On the cognitive effects of learning computer programming. *New Ideas in Psychology*, *2*(2), 137–168. https://doi.org/10.1016/0732-118X(84)90018-7

Perkovic, L., Settle, A., Hwang, S., & Jones, J. (2010). A Framework for Computational Thinking across the Curriculum. *Proceedings of the Fifteenth Annual Conference on Innovation and Technology in Computer Science Education (ITiCS '10)*, 123–127. https://doi.org/10.1145/1822090.1822126

Prat, C. S., Madhyastha, T. M., Mottarella, M. J., & Kuo, C. H. (2020). Relating Natural Language Aptitude to Individual Differences in Learning Programming Languages. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-60661-8

Relkin, E., de Ruiter, L., & Bers, M. U. (2020). TechCheck: Development and Validation of an Unplugged Assessment of Computational Thinking in Early Childhood Education. *Journal of Science Education and Technology*, *29*(4), 482–498. https://doi.org/10.1007/s10956-020-09831-x

Robertson, J., Gray, S., Toye, M., & Booth, J. (2020). The Relationship between Executive Functions and Computational Thinking. *International Journal of Computer Science Education in Schools*, *3*(4), 47–58. https://doi.org/10.21585/ijcses.v3i4.76

Román-González, M. (2015). Computational Thinking Test : Design Guidelines and Content Validation. *Proceedings of EDULEARN15 Conference*, *July*, 2436–2444. https://doi.org/10.13140/RG.2.1.4203.4329

Román-González, M. (2016). *Codigoalfabetización y pensamiento computacional en educación primaria y secundaria: validación de un instrumento y evaluación de programas [Code-literacy and Computational Thinking in Primary and Secondary Education:...]*. 720. http://e-spacio.uned.es/fez/view/tesisuned:Educacion-Mroman

Román-González, M., Moreno-León, J., & Robles, G. (2017). Complementary Tools for Computational Thinking Assessment. *International Conference on Computational Thinking Education 2017*, *July*.

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

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

Rothenbusch, S., Zettler, I., Voss, T., Losch, T., & Trautwein, U. (2016). Exploring reference group effects on teachers' nominations of gifted students. *Journal of Educational Psychology*, *108*(6), 883–897. https://doi.org/10.1037/edu0000085

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

Scherer, R., Siddiq, F., & Viveros, B. S. (2018). Technology and the Mind. *Proceedings of the Technology, Mind, and Society on ZZZ - TechMindSociety '18*, *April*, 1–1. https://doi.org/10.1145/3183654.3183658

Seiter, L., & Foreman, B. (2013). Modeling the learning progressions of computational thinking of primary grade students. *Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research - ICER '13*, 59. https://doi.org/10.1145/2493394.2493403

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

Settle, A., Franke, B., Hansen, R., Spaltro, F., Jurisson, C., Rennert-May, C., & Wildeman, B. (2012). *Infusing computational thinking into the middle- and high-school curriculum*. 22. https://doi.org/10.1145/2325296.2325306

Settle, A., Goldberg, D. S., & Barr, V. (2013). *Beyond computer science*. *July*, 311. https://doi.org/10.1145/2462476.2462511

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

Sneider, C., Stephenson, C., Schafer, B., & Flick, L. (2014). Computational Thinking in High School Science Classrooms: Exploring the Science "Framework" and "NGSS." *Science Teacher*, *81*(5), 53–59. https://www.learntechlib.org/p/155904

Sullivan, A., Kazakoff, E. R., & Bers, M. U. (2013). The wheels on the bot go round and round: Robotics curriculum in pre-kindergarten. *Journal of Information Technology Education*, *12*, 203–219. http://www.jite.org/documents/Vol12/JITEv12IIPp203-219Sullivan1257.pdf

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

Thorndike, R. L., & Hagen, E. P. (1971). *Cognitive Abilities Test.* Houghton-Mifflin.

Thurstone, L. L. (1939). PRIMARY MENTAL ABILITIES: PSYCHOMETRIC MONOGRAPHS No. 1. *British Journal of Educational Psychology*, *9*(3), 270–275. https://doi.org/10.1111/j.2044-8279.1939.tb03214.x

Tsarava, K., Leifheit, L., Ninaus, M., Román-González, M., Butz, M. V., Golle, J., Trautwein, U., & Moeller, K. (2019). Cognitive Correlates of Computational Thinking. *Proceedings of the 14th Workshop in Primary and Secondary Computing Education on - WiPSCE'19*, *October*, 1–9. https://doi.org/10.1145/3361721.3361729

Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2013). Challenges to learning and schooling in the digital networked world of the 21st century. *Journal of Computer Assisted Learning*, *29*(5), 403–413. https://doi.org/10.1111/jcal.12029

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

Weintrop, D., Beheshti, E., Horn, M. S., Orton, K., Trouille, L., Jona, K., & Wilensky, U. (2014). Interactive Assessment Tools for Computational Thinking in High School STEM Classrooms. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, *136 LNICST*, 22–25. https://doi.org/10.1007/978-3-319-08189-2_3

Weintrop, D., & Wilensky, U. (2015). Using commutative assessments to compare conceptual understanding in blocks-based and text-based programs. *ICER 2015 - Proceedings of the 2015 ACM Conference on International Computing Education Research*, *August*, 101–110. https://doi.org/10.1145/2787622.2787721

Weiß, R. H. (2006). *CFT 20-R: grundintelligenztest skala 2-revision.* Hogrefe.

Werner, L., Denner, J., & Campe, S. (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. https://doi.org/10.1145/2157136.2157200

Werner, L., Denner, J., & Campe, S. (2015). Children Programming Games. *ACM Transactions on Computing Education*, *14*(4), 1–22. https://doi.org/10.1145/2677091

Werner, M. (2020). *Computational Thinking in Beziehung zu seinen verwandten psychologischen Konstrukten*. University of Tübingen.

Wiebe, E., Mott, B. W., London, J., Boyer, K. E., Aksit, O., & Lester, J. C. (2019). Development of a lean computational thinking abilities assessment for middle grades students. *SIGCSE 2019 - Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 456–461. https://doi.org/10.1145/3287324.3287390

Wing, J. M. (2006). Computational Thinking. *Theoretical Computer Science*, *49*(3), 33–35. https://doi.org/https://www.cs.cmu.edu/~15110-s13/Wing06-ct.pdf

Zapata-Cáceres, M., Martín-Barroso, E., & Román-González, M. (2020). Computational Thinking Test for Beginners: Design and Content Validation. *2020 IEEE Global Engineering Education Conference (EDUCON)*, 1905–1914. https://doi.org/10.1109/EDUCON45650.2020.91253688

Evaluation of a Computational Thinking Intervention for Elementary School Children: A Randomized Controlled Field Trial

Katerina Tsarava, Luzia Leifheit, Manuel Ninaus, Jessika Golle, Ulrich Trautwein, Korbinian Moeller

Abstract

Computational thinking has been increasingly recognized as a fundamental 21st skill to be fostered early on in education. Accordingly, many governmental, scientific, and professional initiatives have supported developing and implementing educational activities for fostering computational thinking in formal and informal educational settings. The definition and assessment approaches of computational thinking vary and are still undergoing investigation. Consequently, empirical evaluation of the available proposed educational materials is restrained. As part of a larger project, in this work, we present the evaluation of a computational thinking course, which is based on a proposed curriculum we designed for 3rd and 4th graders, utilizing results of our former studies on the cognitive definition of CT and its assessment in primary school. The proposed curricula aims at fostering CT by introducing in unplugged modality various CT concepts, which later on are transferred to different programming environments and are applied in the context of different STEM topics. In a randomized controlled trial with a waiting list control group and pre-/post-test design, we investigated the CT course's effectiveness. Results on a sample of 158 3rd and 4th graders indicated that the intervention had positive effects on students' CT abilities. These results substantiate the results of a pilot study on the course in the field.

Keywords: computational thinking, cognitive skills, computing education research, randomized controlled trials

1 Introduction

Computational Thinking (CT) has gained increasing research attention over the last few years (for a research trends review, see Tang, 2019). As the cognitive underpinning of programming skills and, more generally, as a problem-solving ability, CT has inspired many efforts of integrating it into curricula of different STEM topics across the formal educational levels (e.g., Lockwood & Mooney, 2017; Moreno-Leon et al., 2018). CT has been coined a fundamental cognitive competence to be acquired early on in education, comparable to literacy and numeracy (Wing, 2006; Yadav et al., 2014). In this regard, recently, CT has been for the first time considered for assessment along with other crucial competencies like reading, mathematics, and science skills by the Programme for International Student Assessment (PISA[13]) of OECD[14].

CT has been considered a 21st-century-skill important for students of the present to be prepared for the demands of the increasingly digitized future (D. Barr et al., 2011; Settle et al., 2013; Voogt et al., 2013; Yadav et al., 2011). There are multiple definitions of CT as it was first popularized by Wing (2006) as a fundamental skill that builds upon concepts of computer science (CS), "complements and combines mathematical and engineering thinking" (Wing, 2006, p. 35), but is relevant for everyone beyond computer scientists. There have been numerous efforts to elaborate on a meaningful definition and conceptualize the term (Garcia-Peñalvo, 2016; Grover & Pea, 2013; Lockwood & Mooney, 2017; Yaşar, 2018b). For this work, we consider as a working definition and conceptualization of CT, the interpretation of CT resulting from a literature review by Shute et al. (2017) which describes CT as "the conceptual foundation required to solve problems effectively and efficiently (i.e., algorithmically, with or without the assistance of computers) with solutions that are reusable in different contexts" (Shute et al., 2017, p. 142).

Interpreting CT as a cognitive ability rather than a practical skill (or as a "fundamental, not rote skill"; Wing, 2006, p. 35), we expect its applicability to range across a variety of contexts, not limited to CS-related topics (Armoni, 2016; Settle et al., 2013). This distinction highlights the difference of CT from computer programming or coding, which are more practical and, to a lesser extend, theoretical skills usually referenced as closely related or complementary to CT.

[13] Programme for International Student Assessment: http://www.oecd.org/pisa/
[14] Organisation for Economic Co-operation and Development: http://www.oecd.org/

Though basic CT competencies are required for effective coding and programming, CT as a universal problem-solving approach is associated with developing cognitive skills in a variety of contexts as well (Moreno-Leon et al., 2018; Yaşar, 2018a). In this vein, CT has been integrated into the curricula both as a standalone learning subject (e.g., Chiprianov & Gallon, 2016) and as an interdisciplinary field within different STEM subjects (e.g., Weintrop et al., 2016).

In the following, we first describe the state of CT curricula worldwide, as offered by research, professional and governmental initiatives, as well as popular approaches in teaching CT. Then we present our curricula design approach and the procedure of evaluating our approach in real teaching conditions with a randomized field trial.

1.1 CT curricula worldwide

Despite the different approaches in defining CT, its educational value is broadly argued (Qualls & Sherrell, 2010). Therefore, during the last decade, there were attempts to systematically integrate CT in school and university curricula, supported scientifically by numerous national governments and academic institutions worldwide. Thus, there has been a transparent shift in ICT literacy curricula, from teaching specific skills (e.g., how to create slides for a presentation) on specific tools (e.g., a particular office software) to inspiring thinking on how things work (e.g., how software functions) and potentially foster the creation of digital technology (Curzon et al., 2014). In other words, CT nowadays provides the pathway for students to become potential future creators of technology (or prosumers), and not just consumers of technology (for a discussion on policies towards this direction, see Williamson, 2016).

In this sense, several countries have integrated CT experimentally or officially in their school curricula (e.g., for the UK, see Brown et al., 2014; for France, see Chiprianov & Gallon, 2016; for North Macedonia, see: Jovanov et al., 2016; for Australia, see: Falkner et al., 2014). Along with governmental policies, many professional associations and non-governmental, research, or academic institutions and initiatives have designed and made available curricula to introduce and foster CT at different educational levels. Such a professional association that promotes CT is the Computer Science Teachers Association (CSTA[15]) in the US, which provides

[15] https://www.csteachers.org

the professional CS teachers' community with the most current standards on teaching CS and CT in K-12[16]. A similar multi-background initiative offering teaching resources and qualification seminars to CS teachers is the Computing at School (CAS[17]) initiative in the UK, providing online curricular resources via the Barefoot Computing[18] initiative.

Independently of the CT curricula sources worldwide, an apparent factor of curricula categorization is the target age group. There are curricula available for students from preschool to students at the high school and university level. Curricula addressed to the K-2 are rarer; however, such an example is the San Francisco Unified School District (SFUSD) initiative's colour-coded curricula. Experimental research approaches of CT curricula for early childhood have been reported significantly effective in various studies so far (e.g., Bers et al., 2014; Sullivan et al., 2013). Complete curricula addressed to different school levels of K-12 are offered by the initiative of code.org[19]. CT curricula could also be divided into curricula dedicated to fostering CT as part of Computing Education and curricula integrating CT as a problem-solving approach into other STEM topics beyond CS (V. Barr & Stephenson, 2011; Dierbach et al., 2011). The effectiveness of such curricula has also been explored in various studies (Aggarwal et al., 2017; Duncan & Bell, 2015; Settle et al., 2012; Tran, 2019; Van Dyne & Braun, 2014).

In the population of primary school students, which is of particular interest to this work, some effective curricular interventions have been reported. Rodríguez-Martínez et al. (2020) investigated the effectiveness of a *Scratch*[20] curriculum for learning mathematics and fostering CT in 6th graders. Results indicated no significant differences on CT performance gain between the intervention and the control group. In another study, Brackmann et al. (2017) aimed at improving students' CT skills through unplugged activities (i.e., pen-and-paper activities fostering algorithms, decomposition, pattern recognition, and abstraction) in a sample of 5th and 6th graders and found statistically significant effects in favour of the intervention group. Furthermore, Rose et al. (2019) compared the effectiveness of an educational block-based game intervention (namely *Pirate Plunder*) with a *Scratch*

[16] https://www.csteachers.org/page/about-csta-s-k-12-nbsp-standards
[17] https://www.computingatschool.org.uk
[18] https://www.barefootcomputing.org/curriculum
[19] https://studio.code.org/courses?view=teacher
[20] https://scratch.mit.edu/

intervention (programming control group) and a spreadsheets intervention (non-programming control group) delivered to 10 to 12 years old students. Results indicated a significant effect on students' CT performance for the game intervention group over the non-programming control group.

Moreover, Chiazzese et al. (2019) evaluated a robotics curriculum to foster CT skills in a sample of 3rd and 4th graders. This study's results indicated significant positive effects on CT performance for the intervention group compared to the effects on the control group. In another study, with a younger sample of 9 to 12 years old students, Perez-Marin et al., 2018 found that an intervention utilizing the *Scratch* programming environment and a metaphors' methodology led to significant improvements of students' CT knowledge (Pérez-Marín et al., 2018).

Despite these various studies on effectiveness, a recent review (McGill & Decker, 2020) indicates that less than half of the quantitative studies on K-12 computing education report relevant statistical information like effect sizes, confidence intervals, and levels. This brief overview of evidence on the effectiveness of CT interventions substantiates this claim as only two studies (i.e., Brackmann et al., 2017; Rodríguez-Martínez et al., 2020) reported explicitly and in detail important statistical information. Consequently, the effectiveness of computing and CT curricula still lacks sufficient empirical results.

1.1 CT curricula design

There have been several frameworks for designing CT curricula and implementing them as a broader cognitive ability in a variety of courses, within different contexts, and across educational levels (for an initial framework addressed to university-level students, see Perković et al., 2010; for a framework applied across compulsory education, see Curzon et al., 2014). Though the importance of CT is broadly accepted, and its curricula integration is rapid and continuous, there are still critical open framework questions seeking answers (Chiprianov & Gallon, 2016; Yaşar, 2018a). The *definition* of CT, the *assessment tools* (i.e., *Assessment*) for measuring CT, and its *cognitive aspects (*i.e., *Cognition)*, as well as the appropriate *age* of introducing CT to students (e.g., see Chiprianov & Gallon, 2016), the *context* and *modality* of the materials (e.g., see Brackmann et al., 2017; Wang et al., 2019), the *concepts' interdisciplinarity* (e.g., see Lockwood & Mooney, 2017; Settle et al., 2012), and the *teachers'* *qualifications* on delivering related interventions (e.g., see Angeli et al., 2016; Sentance & Csizmadia, 2017; Yadav et al., 2014) are still not extensively investigated. In previous work (Tsarava et al., *under review*), we aimed to shed light empirically on the first three of these open topics by investigating the cognitive correlates of CT (*Cognition*), by suggesting a reliable CT assessment tool for primary school students (*Assessment*) and therefore supplement the existing cognitive definition of the CT construct (*Definition*).

For providing empirical answers to more open questions about CT, we developed a CT course for primary school students (*Age*), which introduces basic CT processes and coding skills (*Context*), applied in various STEM contexts, like Math, Biology, Technology, etc. (*Context*), in a playful, multimodal way (i.e., unplugged/plugged-in activities, board games, and playful digital activities; haptic/visual demonstrations; *Modality*). We intended to provide a broader perspective on the applicability of CT, not only related to CS topics but also applied in activities of other STEM domains (like math applications, biology animations, robot simulations, and game production; *Concepts' Interdisciplinarity*), and subsequently promote the importance of understanding basic CT concepts.

Taking into consideration common concerns about the introduction of coding already in primary school (see Garcia-Peñalvo, 2016), we designed a course that introduces and fosters broader CT concepts in the context of coding (*Context*), and we aimed for a low threshold introduction to these concepts in a playful and unplugged (i.e., no computers devices involved) way (*Modality; Age*). The course was constructed on the principles of *learning-by-doing* as they derive from Papert's constructionism (Papert & Harel, 1991), and they are interpreted within the context of CS (Sentance & Csizmadia, 2017).

The CT course focuses on the following: i) material design that stimulates game-based learning and promotes learning-by-doing, ii) content that fosters the cognitive processes of CT and not only the practical coding skills, iii) unplugged low-threshold introduction to CT and coding concept that gradually transfers into plugged-in environments (for the initial conceptualization of the course, see Tsarava et al., 2017; for an updated structure of the course, see Tsarava et al., 2019).

For a broader applicability of the developed course and to avoid any platform dependencies, we suggest that at an early age, learning to code should be detached from a specific programming language (*Age*). The programming language trends, highly dependent on the technology market developments, change rapidly. For that reason, any attachment to a specific language at that early stage would not necessarily benefit students later on as adults when the programming language trends would possibly have changed. Thereby, we initially introduced CT and coding concepts unplugged by developing a series of life-size unplugged games (i.e., *Crabs & Turtles: A Series of Computational Adventures*[21]; Tsarava et al., 2018; Tsarava, Moeller, et al., 2019) that offer an age-appropriate, haptic and low-threshold introduction to CT and coding concepts (*Age; Modality*). The same concepts are later on in the course introduced in different plugged-in visual-programming environments (like *Scratch*, *Scratch for Arduino*[22], and *Open Roberta Lab*[23]), specially designed for young novice programmers.

Along with the course content, we developed a detailed course manual, which provides course instructors with the essential methodological background on our design decisions and methods implemented in the course, as well as detailed lesson plans on how to implement each course unit step-by-step (*Teachers' qualification*). This manual offers both the theoretical background needed for understanding CT and the teaching methods used (the cognitive ladder designed for supporting the scaffolding while learning about CT), along with practical knowledge on the content (CT concepts and plugged-in/unplugged materials), the activities, and the assessments used.

1.2 Aim of the study

In a pilot study (Tsarava, Leifheit, et al., 2019), we delivered the proposed intervention to a small number of children as part of an extracurricular enrichment program for elementary school children. Moving one step forward, in the current study, we evaluated the effectiveness of the intervention in a randomized field trial employing a waiting list control group design and involving 197 children at 16 different sites. The basic parameters of the study (e.g., content and duration of the intervention) remained the same. Importantly, however, the

[21] *Crabs & Turtles* are available as an open educational resource (OER) via the OER repository of the University library of Tübingen. The OER is available at http://hdl.handle.net/10900.3/OER_MDCKSMXP.
[22] http://s4a.cat/
[23] https://lab.open-roberta.org/

intervention was delivered by a group of trained instructors in a natural setting. The details of the current study, including the intervention, are described in detail in the following section.

Based on the findings of the pilot study, we pursued the following hypothesis: We expected that children participating in the CT course would improve their CT performance more strongly as reflected by a CT assessment compared to the children of the control group (Hypothesis).

2 Method

2.1 Research design

A randomized controlled field trial with pre- and post-intervention measurements and a waiting list control group was employed to assess the CT intervention's effectiveness. Participants were assigned randomly to either the experimental or the control group. Students' allocation to either control or experimental group was done after pre-testing to ensure unbiased performance at the pre-test (see Figure 1). The study was approved by the Ethics Committee of the Leibniz-Institut für Wissensmedien, Tübingen.

Randomization was done based on random computer-generated numbers and took place after the pre-test procedure was over. The pre-test instructors, who were either the developers of the course or trained student assistants, conducted the randomization, and the procedure was blind to the course instructors. After that, students were informed about the group they belonged to, and course instructors were informed about each group's students. All children enrolled in the experimental group took the intervention weekly for ten consecutive weeks (except for interfering school vacation weeks) after the pre-test session. The post-test took place one week after the end of the intervention, and students of both the experimental and the control group were invited to participate. Afterwards, at a prescheduled appointment already known to the students during their initial registration to the course, the students of the control group were offered the intervention as a block course, organized in fewer but longer course sessions.

The CT course was offered as an intervention in after-school academies. The intervention consisted of ten 90-minutes sessions and was designed for groups of five to 10 3rd and 4th graders. This study considered courses held at 16 local sites of the Hector Children's Academy Program (HCAP) during the winter semester of 2018/2019. The HCAP is an extracurricular program for elementary school students implemented at 66 different sites across the state of

Baden-Württemberg, Germany. Children nominated by their teachers can participate in the HCAP enrichment program (Rothenbusch et al., 2016), which offers a variety of courses taking place in small group after-school classes or during weekends.

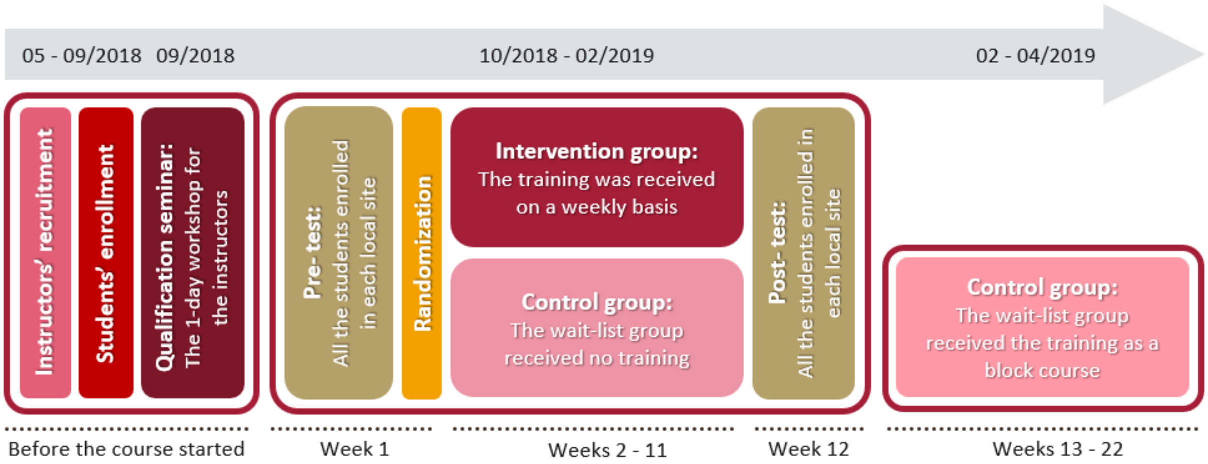

Figure 1. Timeline of study procedures.

The courses at the HCAP are taught by experienced instructors (teachers and professionals of different fields). To recruit instructors for the present intervention, we sent informational material about the course content and aims, as well as the experimental design of the study, to all local HCAP sites, five months before the beginning of the study. Twenty-eight instructors registered for participating in the study; however, only the first twenty were considered for participation due to organizational matters.

2.2 Intervention: The CT course

The CT course aimed at fostering CT skills of 3rd and 4th graders conceptualized for classes of five to ten children. The course introduced and trained CT concepts as identified within coding and problem-solving activities by Brennan & Resnick (2012). More specifically, the concepts considered are the ones of sequences, loops, patterns, events, conditionals, events, operators, data (i.e., variables and constants), algorithms, and simulations.

The instructional design of the course supported the introduction of each concept in a multimodal way, integrated into both non-programming (i.e., unplugged) and programming (i.e., plugged-in) activities. Students are first introduced to the respective CT concepts by engaging in playful unplugged activities. They are then guided to modify elements of existing plugged-in activities, and subsequently, they are motivated to create their own functional applications (Tsarava et al., 2017; for a detailed course overview, see Figure 2). The course design reinforced the applicability of the CT concepts within different STEM disciplines. Each of the ten-course sessions of the intervention contextualized in a particular STEM discipline providing activities related to this specific topic. The aim was to broaden students' perspectives on CT and programming as essential skills for problem-solving in everyday life, within various contexts beyond CS. In the course, students start by programming simple games, and then simple and more complex simulations inspired by topics inspired from biology, mathematics, science, and engineering.

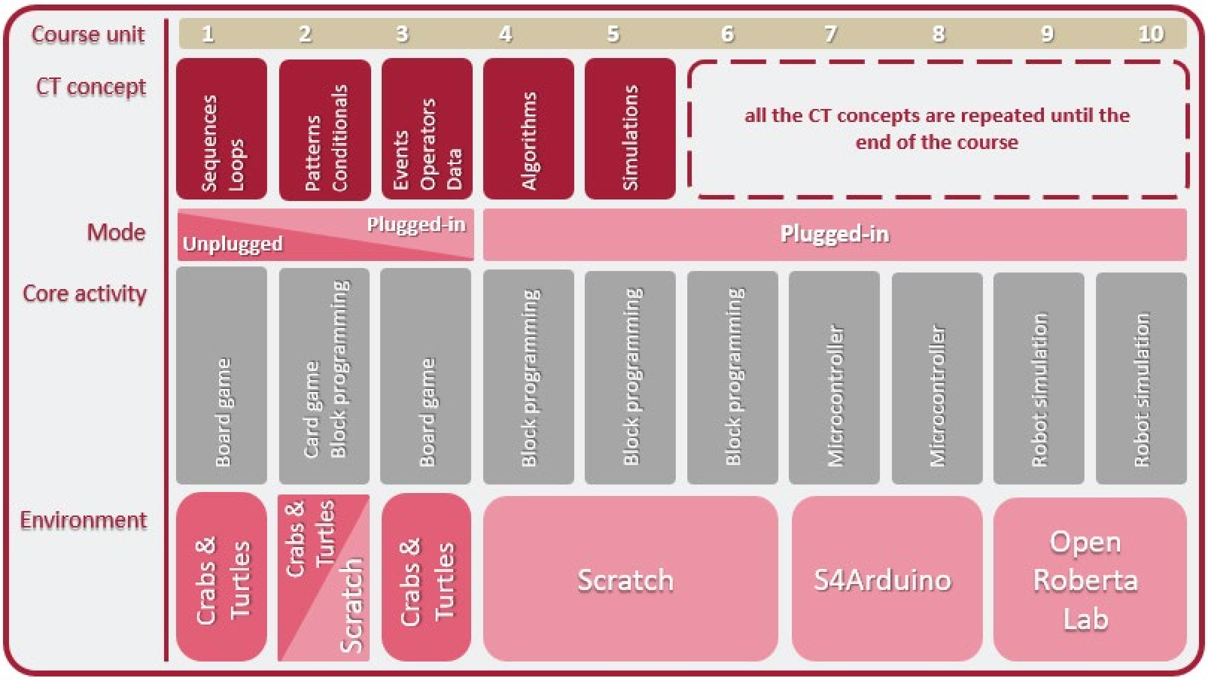

Figure 2. Course overview.

In course sessions 1, 2, and 3, students were playfully introduced to basic CT concepts, making minimal use of technology. In particular, in course unit 1, students played "The Hunt" game of the CT game series "Crabs & Turtles" in teams of two (for a description of the game, see Tsarava et al., 2018). In this life-size board game, students got introduced to the concepts of sequences and loops in an unplugged mode and learned to build their sequences of commands in order to achieve their strategical moves towards collecting treasures on the 2D board. In course unit 2, the concept of patterns was introduced along with an implicit reference to conditionals. The students first played the card game "Patterns" from the CT game series "Crabs & Turtles", where they had to recognize, interpret, and conditionally match pattern cards with each other. Then, students got introduced to the block-programming environment of *Scratch*. While working in teams of 2 per machine (laptop or

PC), they got to know the essential elements of the block programming environment and how to build short algorithms. Course unit 3 was implemented unplugged, and students, again in teams of 2, engaged in “The Race” game of “Crabs & Turtles” (for details on the game, see Tsarava et al., 2018). The CT concepts introduced in this session were events, operators, and data (i.e., variables and constants). Students had to correctly solve mathematical riddles, which integrated variables with changing values and constants, in order to first arrive at the finish line.

In course sessions 4, 5, and 6, students worked in the visual block-programming environment of *Scratch*. In the activities of these sessions, students applied the conceptual knowledge acquired while playing the unplugged games of “Crabs & Turtles” to activities of gradually increased difficulty programmed with *Scratch*. First, they started by programming two short games. They learned how to create, coordinate objects on the screen and program them, and assign keyboard reactions to their programs. Additionally, they implemented conditional logic for programming the winning and losing conditions of their games. Second, they programmed two simulations inspired by biology topics. The simulations demanded to conditionally present or hide objects from the main screen and repeatedly implement them using loop constructs. Third, students programmed a complex simulation inspired by math concepts. In this activity, students had to handle constants and variables embedded in conditional and loop constructs that interact with different events.

In course units 7 and 8, students worked with the Arduino hardware platform, programmed in *Scratch* (S4A). In the activities of these units, students familiarized themselves with the CT concepts already acquired in the previous units, programmed in *Scratch*, and applied on a physical output equipped with physical sensors. The activities include interaction with light sensors, led lights, and buttons, and implement concepts like a Morse code machine, a light-controller, and a traffic light. The main objective of all activities was to show the interaction of programming in a computer environment with physical objects, and therefore the broad applicability of CT concepts.

In the last two course units of the intervention, 9 and 10, students exercised the CT concepts independently by programming a robot simulation in the interactive programming environment Open Roberta Lab. In the activities of these course units, students made use of all the CT concepts repeatedly introduced throughout the course, used their knowledge on sensors' functionality already introduced in the sessions with Arduino, and programmed their robots in a simulated environment using visual block programming. The robot simulation activities were less guided and more open for students to set specific goals, define the particular problems, and individually design the required strategies to solve them.

2.3 Sample

Data were collected from 197 students (47 girls, age: *M* = 9.13, *SD* = 0.61; 150 boys; age: *M* = 9.20, *SD* = 0.70) in Grade 4 (*N* = 111; girls =24, boys = 87), in Grade 3 (*N* = 85; girls = 23, boys = 62), and 1 in Grade 2. Children in both the intervention and the control group were assigned to 29 different courses (16 intervention; 13 control). The difference in the number of intervention and control groups was due to randomization conditions. The intervention group consisted of 103 students, 24 girls (*M* = 8.98, *SD* = 0.60) and 79 boys (*M* = 9.24, *SD* = 0.67). The control group consisted of 94 students, 23 girls (*M* = 9.28, *SD* = 0.58) and 71 boys (*M* = 9.17, *SD* = 0.73). In case a local HCAP site had more than five students registered for the CT course, then students would randomly be assigned in two groups (one intervention and one control group). When only five or fewer students registered for the CT course, then there was only one group created in this HCAP site, which was randomly assigned to be either a control or intervention group. Parents provided written informed consent before the start of the study.

2.4 Implementation

Instructors of the CT course got qualified in a one-day-long qualification seminar where they were presented with the course aims, a comprehensive teaching manual, and evaluation procedure. During the qualification seminar, instructors were provided with all teaching materials, including a scripted course manual, printable activity- and assessment-sheets, a games box, and an Arduino kit. The qualification seminar took place twice in September 2018.

The course manual offered a description of the general theoretical background of computational thinking, of the teaching methods followed in the course, provided information for each of the course units, the goals of each unit, the introductory exercises, and the assessment sheets for each unit, along with specific time frames for the better time planning.

2.5 Measures

All the performance measures were administered twice, one week before the beginning and one week after the end of the course. The measures included several cognitive assessments and a CT assessment. Assessment instruments were the following:

i. *Mathematical skills:* We used five subtests of the Heidelberger Rechentest zur Erfasssung mathematischer Basiskompetenzen im Grundschulalter (HRT 1-4; Haffner et al., 2005) to assess students' mathematical skills in speeded subtests. The respective subscales used were the *Writing Speed* (time-limit of 30 seconds), the *Addition* (2 minutes), the *Subtraction* (2 minutes), the *Multiplication* (fact retrieval; 2 minutes), and the *Problem Completion* (2 minutes). Each mathematical subtest consisted of 40 arithmetic problems, and students had to complete as many tasks as possible within the given time limit.

ii. *Non-verbal visuospatial reasoning:* We used two subtests of the Culture Fair Intelligence Scale (CFT 20-R; Weiß, 2006) to assess students' non-verbal visuospatial skills. The subtests used were the *Continuing Series* (Subtest 1; time-limit of 4 minutes) and the *Matrices* (Subtest 3; 3 minutes). The participants had in total 7 minutes to complete both subtests.

iii. *Verbal reasoning:* We used the subtest *V1* (Form A) of the Kognitiver Fähigkeitstest (KFT 4-12+R; Heller & Perleth, 2000) in order to assess verbal reasoning abilities. Participants had 7 minutes to complete as many items as possible.

iv. *Computational thinking:* To assess CT abilities, we used the Abbreviated Computational Thinking test (*Abbreviated CTt;* Tsarava et al., 2020 *under review*). The test consists of 21 items of the Computational Thinking test (*CTt*; Román-González et al., 2017), translated in German. The seven different CT concepts addressed by the test are i. basic directions and sequences, ii. loops implemented with repeat-times commands, iii. loops implemented with repeat-until commands, iv. simple conditionals implemented with if commands, v. complex conditionals, implemented with if/else commands, vi. while conditionals, and vii. simple functions. The participants were given a time-limit of 20 minutes to complete the test.

2.6 Statistical analyses

From each of the aforementioned tests, the following variables were considered in the analyses:

i. WRSpre and WRSpost: number of correct responses in the *Writing Speed* subtest of the HRT 1-4, at pre- and post-test assessment.
ii. CALCpre and CALCpost: mean of correct responses in the *Addition* and *Subtraction* subtests.
iii. MULpre and MULpost: number of correct responses in the *Multiplication* subtest.
iv. COMpre and COMpost: number of correct responses in the *Problem Completion* subtest.
v. CFTpre and CFTpost: mean of correct responses in the *Continuing Series* and *Matrices* subtests of the CFT.
vi. KFTpre and KFTpost: number of correct responses in the verbal reasoning subtest of the KFT.
vii. CTTpre and CTTpost: number of correct responses in the *abbreviated CTt*.

To investigate the effectiveness of the CT course, we conducted a repeated-measures ANCOVA analysis discerning the within-participant factor measurement time (pre- vs post-test) and the between-participant factor group (intervention vs control group). This, along with the descriptive analyses, were run in SPSS version 26 (IBM Coorporation 1989-2019).

Results of a correlational analysis of CT with other cognitive abilities conducted on pre-test performance are shown in Table 1 (for more information, see Tsarava et al., 2020 -under review-). Based on the observed associations of CT with other cognitive skills assessed at pre-test, we considered performance at pre-test of i. the *abbreviated CTt* (CTTpre), ii. the *Problem Completion* (COMpre), iii. the *Continuing Series* and *Matrices* (CFTpre), and iv. the verbal reasoning subtest (KFTpre) as covariates in the ANCOVA. Our selection includes performance scores with a Pearson's r of at least $r = .3$ with CT performance.

Table 1. Correlations (Pearson's r) between CT performance and other cognitive abilities scores.

	WRSpre	CALCpre	MULpre	COMpre	CFTpre	KFTpre
CTTpre	-.023	.232**	.224**	**.333****	**.346****	**.388****

**. Correlation is significant at the 0.01 level (2-tailed).

Participants were randomly assigned to the experimental or the control conditions. Nevertheless, to ensure that the two groups were comparable at pre-test, we evaluated baseline performance at pre-test by running two-sample *t*-tests for all control variables (CTTpre, COMpre, CFTpre, KFTpre).

3 Results

3.1 Descriptive statistics

From the 197 students participating in the study, data of 158 students that completed both pre- and post-test was considered in the ANCOVA. The detailed characteristics of this sample are presented in Table 2.

Table 2. Description of the sample considered in the regression analysis.

Group		Age	Grade 2	Grade 3	Grade4
Intervention group	Girls *N* = 20	*M* = 9.05 *SD* = 0.58	*N* =0	*N* =11	*N* =9
	Boys *N* = 63	*M* = 9.21 *SD* = 0.69	*N* =1	*N* =24	*N* =38
Control group	Girls *N* = 15	*M* = 9.49 *SD* = 0.51	*N* =0	*N* =4	*N* =11
	Boys *N* = 60	*M* = 9.16 *SD* = 0.76	*N* =0	*N* =27	*N* =33

3.2 Course effects

To evaluate the effects of the intervention, an ANCOVA was conducted controlling for influences of CTTpre, EGpre, KFTpre, and CFTpre (see Table 3 for descriptives of the variables). As dependent variable, we considered the learning gain computed by the difference of the *CTt* performance scores from pre- to post-test [CTT = CTTpost – CTTpre]. Testing our directed hypothesis of a more pronounced improvement in CT skills for those children who completed the course was substantiated by the results [$t(152) = 1.88$, $p = .031$, tested one-sided, ANCOVA *F*-test: $F(1, 152) = 3.548$, $p = .062$]. The effect size, according to Morris (2008), was $d_{ppc2}=0.218$, which indicates a rather small effect (Cohen, 1988).

4 Discussion

CT, as a 21st-century-skill, has attracted significant research attention in recent years (e.g., Tang, 2019). A plethora of initiatives worldwide have suggested and worked towards the integration of CT in formal and non-formal educational settings by providing guidelines and teaching materials for different age groups, either as short standalone interventions or as part of integrative curricula. Consequently, this increasing implementation of CT activities requires corresponding empirical studies to evaluate the materials and methods used.

Table 3. Means and standard deviations of all variables for each measurement point and group.

	Pre-test				Post-test			
Group	Intervention		Control		Intervention		Control	
Variables	M	SD	M	SD	M	SD	M	SD
Dependent								
Computational thinking (CTT)	11,10	3,38	11,37	3,36	14,30	5,70	13,83	3,58
Control								
Mathematical skills-Number Completion (EG)	14,93	5,70	16,15	6,14	18,67	5,81	19,33	6,71
Verbal reasoning (KFT)	15,16	3,45	15,28	4,02	16,73	3,63	17,35	3,89
Non-verbal visuospatial reasoning (CFT)	9,83	2,01	10,13	2,01	10,67	1,80	10,75	1,74

Despite the constant and rapid developments for integrating CT in educational settings, there are several open questions seeking answers regarding the *definition*, the *assessment*, and the *cognition* of CT, along with the appropriate *age*, *context* and *modality* of introducing it, and the *teachers'* required *qualifications* for doing so (Brackmann et al., 2017; Chiprianov & Gallon, 2016; Lockwood & Mooney, 2017; Sentance & Csizmadia, 2017; Settle et al., 2012; Wang et al., 2019; Yadav et al., 2014; Yaşar, 2018b). In this work, we present the continuation of prior work we conducted, aiming to answer these open questions empirically.

In previous work (Tsarava et al., *under review*) we focused on the first three open questions (*Definition*, *Assessment*, *Cognition*) by i. conducting a correlational analysis between CT and other cognitive abilities, and ii. suggesting a CT assessment for primary school students. The correlational analysis complemented the current idea of CT as a cognitive skill well integrated with other cognitive skills and, therefore, our understanding of the construct. The CT assessment developed seems to provide a reliable measure for evaluating CT in primary school. In the current study, we evaluated the effectiveness of a CT curriculum we designed in a randomized controlled field trial to provide empirical answers to the questions regarding the appropriate *age*, *context,* and *modality* for introducing CT, and *teachers'* required *qualifications* for delivering CT courses. Notably, the intervention was instructed by trained instructors in real classroom settings.

The CT course we designed was built based on constructivist principles, integrating game-based learning, learning-by-doing, and embodied learning. The course introduced basic CT processes as applied in coding and implemented through activities of various STEM domains. To be age-appropriate, the different CT processes were first introduced unplugged via life-size board games, and later on, trained in different plugged-in activities implemented in visual programming environments.

In the following, we elaborate on the course effects' interpretation before discussing the potential limitations of the current study and future perspective.

4.1 Interpretation of the Effects

The CT course had a significant positive effect on students' CT abilities. Students in the experimental group showed a larger improvement in their CT performance from pre- to post-test than observed for the students in the control group. Our study used a randomized controlled trial design that allows for reliable evidence to interpret the significant beneficial effects of the intervention. Importantly, performance improvements from pre- to post-test were observed for both groups (as to be expected when participants are tested twice on the same test). However, the improvement of the experimental group was significantly larger than the improvement of the control group.

Compared to our pilot study where no control group design was implemented, the current results substantiated evidence on the beneficial effects of the course on participants' CT skills (see Tsarava, Leifheit, et al., 2019) – even when delivered by teachers in a natural educational setting. They, therefore, reflect the effectiveness of the proposed CT course. Nevertheless, the rather small effect size needs to be considered when interpreting the results.

To the best of our knowledge, there are no comparable intervention studies that implement an unplugged/plugged-in approach to foster CT in an extracurricular setting in the age group of 3rd and 4th graders. A recent review of the K-12 Computing Education research in the US indicated that research findings in extracurricular settings after 2016 are not available, and research papers with elementary school populations are significantly more seldom than respective papers involving middle and high school students (Upadhyaya et al., 2020).

Due to the lack of comparable research, a direct comparison of effect sizes with similar studies cannot be performed. Nevertheless, a recent meta-analysis of children's learning outcomes in block-based programming courses found that in the existing literature, only small effect sizes for students' CT improvement are reported (Chiu & Tsuei, 2020). Since 7 out of the 10 units of our CT course consisted of visual block-based programming activities, a distant comparison of our effect size with this meta-analysis results confirms this main finding.

4.2 Limitations

We are aware that two limitations might have influenced the results of the current work. The first is the possibility of self-selection bias of our sample. The students of our sample participated in the course intervention of their will. They were not assigned to the CT course, but they selected for themselves the course from various available courses offered in their academy during the respective school semester. For that reason, our results may not be generalized to the broader educational context. Secondly, our sample's self-selection attribute may have also affected the ratio of girls and boys, with a significantly larger number of boys participating in the CT course.

4.3 Future perspective

Despite the above-described limitations, this study comprehensively investigated the effects of a course on CT development in a sample of primary school students. The intervention was offered by trained teachers (and not by the developers of the course), which allows for the conclusion that the course seems effective in real-class conditions. We hope that future analyses of the data on treatment fidelity questionnaires will shed light on a greater detail of each course unit's feasibility.

ACKNOWLEDGEMENTS

We want to thank the following colleagues for their support in this study; our student research assistants Jana Hofmann, Moritz Werner, Joachim Fritscher, Joshua Schmid, and Daniela Piechnik, our colleague Silke Wortha, as well as the bachelor's students Christian Reiff, Mareike Nutz, Marcel Fröhlich, and Ioana Uhl. Additionally, we gratefully acknowledge the teachers and headmasters of the 16 Hector Children's Academies that participated in the study and the team "Wissenschaftliche Begleitung der Hector Kinderakademien" at the Hector Research Institute of Education Sciences and Psychology, that constantly supported us with research feedback and administrative support. This research was primarily supported by the Hector Foundation II and the LEAD Graduate School and Research Network funded by the Excellence Initiative of the German federal and state governments.

References

Aggarwal, A., Gardner-Mccune, C., & Touretzky, D. S. (2017). Evaluating the effect of using physical manipulatives to foster computational thinking in elementary school. *Proceedings of the Conference on Integrating Technology into Computer Science Education, ITiCSE*, 9–14. https://doi.org/10.1145/3017680.3017791

Angeli, C., Voogt, J., Fluck, A., Webb, M., Cox, M., Malyn-Smith, J., & Zagami, J. (2016). International Forum of Educational Technology & Society A K-6 Computational Thinking Curriculum Framework: Implications for Teacher Knowledge. *Journal of Educational Technology & Society*, *19*(3), 47–57.

Armoni, M. (2016). COMPUTING IN SCHOOLS Computer science, computational thinking, programming, coding. *ACM Inroads*, *7*(4), 24–27. https://doi.org/10.1145/3011071

Barr, D., Harrison, J., & Conery, L. (2011). Computational Thinking: A Digital Age Skill for Everyone. *Learning and Leading with Technology*, *38*(6), 20–23.

Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12. *ACM Inroads*, *2*(1), 48–54. https://doi.org/10.1145/1929887.1929905

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

Brackmann, C. P., Román-González, M., Robles, G., Moreno-León, J., Casali, A., & Barone, D. (2017). *Development of Computational Thinking Skills through Unplugged Activities in Primary School*. 65–72. https://doi.org/10.1145/3137065.3137069

Brennan, K., & Resnick, M. (2012). Using artifact-based interviews to study the development of computational thinking in interactive media design. *Proceedings of the Annual American Educational Research Association Meeting (AERA)*. http://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf

Brown, N. C. C., Sentance, S. U. E., Crick, T. O. M., & Humphreys, S. (2014). Restart: The Resurgence of Computer Science in UK Schools. *ACM Transactions on Computing Education*, *14*(2), 1–22. https://doi.org/10.1145/2602484

Chiazzese, Arrigo, Chifari, Lonati, & Tosto. (2019). Educational Robotics in Primary School: Measuring the Development of Computational Thinking Skills with the Bebras Tasks. *Informatics*, *6*(4), 43. https://doi.org/10.3390/informatics6040043

Chiprianov, V., & Gallon, L. (2016). Introducing Computational Thinking to K-5 in a French Context. *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '16*, 112–117. https://doi.org/10.1145/2899415.2899439

Chiu, J.-I., & Tsuei, M. (2020). *Meta-Analysis of Children's Learning Outcomes in Block-Based Programming Courses* (pp. 259–266). https://doi.org/10.1007/978-3-030-60703-6_33

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers.

Curzon, P., Dorling, M., Selby, C., Woollard, J., & Ng, T. (2014). *Developing computational thinking in the classroom: a framework*. *June*. http://eprints.soton.ac.uk/369594/10/DevelopingComputationalThinkingInTheClassroomaFramework.pdf

Dierbach, C., Hochheiser, H., Collins, S., Jerome, G., Ariza, C., Kelleher, T., Kleinsasser, W., Dehlinger, J., & Kaza, S. (2011). A Model for Piloting Pathways for Computational Thinking in a General Education Curriculum. *Development*, *15*(5), 257–262. https://doi.org/10.1145/1953163.1953243

Duncan, C., & Bell, T. (2015). A pilot computer science and programming course for primary school students. *ACM International Conference Proceeding Series*, *09-11-Nove*, 39–48. https://doi.org/10.1145/2818314.2818328

Falkner, K., Vivian, R., & Falkner, N. (2014). The Australian digital technologies curriculum: Challenge and opportunity. *Conferences in Research and Practice in Information Technology Series*, *148*(January), 3–12.

Garcia-Peñalvo, F. J. (2016). What Computational Thinking Is. *Journal of Information Technology Research*, *9(3)*, *v–vi*(October).

Grover, S., & Pea, R. (2013). Computational Thinking in K-12: A Review of the State of the Field. *Educational Researcher*, *42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

Haffner, J., Baro, K., Parzer, P., & Resch, F. (2005). *HRT1-4: Heidelberger Rechentest; Erfassung mathematischer Basiskompetenzen im Grundschulalter.* Hogrefe.

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision 3.* Beltz Test.

Jovanov, M., Stankov, E., Mihova, M., Ristov, S., & Gusev, M. (2016). Computing as a new compulsory subject in the Macedonian primary schools curriculum. *IEEE Global Engineering Education Conference, EDUCON*, *10-13-Apri*(April), 680–685. https://doi.org/10.1109/EDUCON.2016.7474623

Lockwood, J., & Mooney, A. (2017). *Computational Thinking in Education : Where does it Fit ? A systematic literary review A systematic literary review*. *March*, 1–58.

McGill, M. M., & Decker, A. (2020). A Gap Analysis of Statistical Data Reporting in K-12 Computing Education Research. *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 591–597. https://doi.org/10.1145/3328778.3366842

Moreno-Leon, J., Roman-Gonzalez, M., & Robles, G. (2018). On computational thinking as a universal skill: A review of the latest research on this ability. *2018 IEEE Global Engineering Education Conference (EDUCON)*, 1684–1689. https://doi.org/10.1109/EDUCON.2018.8363437

Morris, S. B. (2008). Estimating Effect Sizes From Pretest-Posttest-Control Group Designs. *Organizational Research Methods*, *11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Papert, S., & Harel, I. (1991). *Constructionism*. Ablex Publishing Corporation.

Pérez-Marín, D., Hijón-Neira, R., Bacelo, A., & Pizarro, C. (2018). Can computational thinking be improved by using a methodology based on metaphors and scratch to teach computer programming to children? *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2018.12.027

Perković, L., Settle, A., Hwang, S., & Jones, J. (2010). A framework for computational thinking across the curriculum. *Proceedings of the Fifteenth Annual Conference on Innovation and Technology in Computer Science Education - ITiCSE ’10*, 123. https://doi.org/10.1145/1822090.1822126

Qualls, J. A., & Sherrell, L. B. (2010). Why computational thinking should be integrated into the curriculum. *Journal of Computing Sciences in Colleges*, *25*(5), 66–71.

Rodríguez-Martínez, J. A., González-Calero, J. A., & Sáez-López, J. M. (2020). Computational thinking and mathematics using Scratch: an experiment with sixth-grade students. *Interactive Learning Environments*, *28*(3), 316–327. https://doi.org/10.1080/10494820.2019.1612448

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

Rose, S. P., Habgood, M. P. J., & Jay, T. (2019). Using Pirate Plunder to Develop Children's Abstraction Skills in Scratch. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, *May*, 1–6. https://doi.org/10.1145/3290607.3312871

Rothenbusch, S., Zettler, I., Voss, T., Losch, T., & Trautwein, U. (2016). Exploring reference group effects on teachers' nominations of gifted students. *Journal of Educational Psychology*, *108*(6), 883–897. https://doi.org/10.1037/edu0000085

Sentance, S., & Csizmadia, A. (2017). Computing in the curriculum: Challenges and strategies from a teacher's perspective. *Education and Information Technologies*, *22*(2), 469–495. https://doi.org/10.1007/s10639-016-9482-0

Settle, A., Franke, B., Hansen, R., Spaltro, F., Jurisson, C., Rennert-May, C., & Wildeman, B. (2012). *Infusing computational thinking into the middle- and high-school curriculum*. 22. https://doi.org/10.1145/2325296.2325306

Settle, A., Goldberg, D. S., & Barr, V. (2013). *Beyond computer science*. *July*, 311. https://doi.org/10.1145/2462476.2462511

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

Sullivan, A., Kazakoff, E. R., & Bers, M. U. (2013). The wheels on the bot go round and round: Robotics curriculum in pre-kindergarten. *Journal of Information Technology Education*, *12*, 203–219. http://www.jite.org/documents/Vol12/JITEv12IIPp203-219Sullivan1257.pdf

Tang, K. (2019). A Content Analysis of Computational Thinking Research : An International Publication Trends and Research Typology. *The Asia-Pacific Education Researcher*. https://doi.org/10.1007/s40299-019-00442-8

Tran, Y. (2019). Computational Thinking Equity in Elementary Classrooms: What Third-Grade Students Know and Can Do. *Journal of Educational Computing Research*, *57*(1), 3–31. https://doi.org/10.1177/0735633117743918

Tsarava, K., Leifheit, L., Ninaus, M., Román-González, M., Butz, M. V., Golle, J., Trautwein, U., & Moeller, K. (2019). Cognitive Correlates of Computational Thinking. *Proceedings of the 14th Workshop in Primary and Secondary Computing Education on - WiPSCE'19*, *October*, 1–9. https://doi.org/10.1145/3361721.3361729

Tsarava, K., Moeller, K., & Ninaus, M. (2018). Training Computational Thinking through board games: The case of Crabs & Turtles. *International Journal of Serious Games*, *5*(2), 25–44. https://doi.org/10.17083/ijsg.v5i2.248

Tsarava, K., Moeller, K., & Ninaus, M. (2019). Board Games for Training Computational Thinking. In M. Gentile, M. Allegra, & H. Söbke (Eds.), *Games and Learning Alliance* (pp. 90–100). Springer International Publishing. https://doi.org/10.1007/978-3-030-11548-7_9

Tsarava, K., Moeller, K., Pinkwart, N., Butz, M. V., Trautwein, U., & Ninaus, M. (2017). Training computational thinking: Game-based unplugged and plugged-in activities in primary school. *Proceedings of the 11th European Conference on Games Based Learning, ECGBL 2017*, *October*, 687–695.

Upadhyaya, B., McGill, M. M., & Decker, A. (2020). A longitudinal analysis of k-12 computing education research in the united states: Implications and recommendations for change. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, 605–611. https://doi.org/10.1145/3328778.3366809

Van Dyne, M., & Braun, J. (2014). Effectiveness of a computational thinking (CS0) course on student analytical skills. *Proceedings of the 45th ACM Technical Symposium on Computer Science Education - SIGCSE '14*, 133–138. https://doi.org/10.1145/2538862.2538956

Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2013). Challenges to learning and schooling in the digital networked world of the 21st century. *Journal of Computer Assisted Learning*, *29*(5), 403–413. https://doi.org/10.1111/jcal.12029

Wang, P., Fessard, G., & Wang, P. (2019). *Are There Differences in Learning Gains When Programming a Tangible Object or a Simulation ? Are There Differences in Learning Gains When Programming a Tangible Object or a Simulation ? July*. https://doi.org/10.1145/3304221.3319747

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

Weiß, R. H. (2006). *CFT 20-R: grundintelligenztest skala 2-revision.* Hogrefe.

Williamson, B. (2016). Political computational thinking: policy networks, digital governance and 'learning to code.' *Critical Policy Studies*, *10*(1), 39–58. https://doi.org/10.1080/19460171.2015.1052003

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

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational Thinking in Elementary and Secondary Teacher Education. *ACM Transactions on Computing Education*, *14*(1), 1–16. https://doi.org/10.1145/2576872

Yadav, A., Zhou, N., Mayfield, C., Hambrusch, S., & Korb, J. T. (2011). Introducing computational thinking in education courses. *Proceedings of the 42nd ACM Technical Symposium on Computer Science Education - SIGCSE '11*, *2*, 465. https://doi.org/10.1145/1953163.1953297

Yaşar, O. (2018a). A new perspective on computational thinking. *Communications of the ACM*, *61*(7), 33–39. https://doi.org/10.1145/3214354

Yaşar, O. (2018b). Computational Thinking, Redefined. In E. Langran & J. Borup (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (Issue June, pp. 72–80). Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/primary/p/182505/

PART III: GENERAL DISCUSSION

8 Summary of Results

In this section, a summary of the six enclosed studies' results will be presented in order of conduct (sections 8.1 to 8.6). The aim of this thesis is a cognitive definition of CT that will allow the development of more appropriate CT assessment tools and, therefore, contribute to evaluating educational materials designed for developing CT. The appropriate evaluation of CT interventions' learning outcomes can lead to more efficient didactical approaches for introducing and fostering CT. This thesis focuses on the elementary school level and aims to complement the CT research conducted in other age groups (e.g., at middle and high-school levels) so far.

8.1 Findings of Study 1: CT curriculum design

In Study 1, a review of the most recent literature was conducted, and based on it, the initial conceptualization of a CT curriculum for elementary school students was developed. In this study, a first break down of relevant CT processes was formulated (i.e., *decomposition*, *algorithms*, *logic*, *patterns*, *evaluation*, *abstraction*, and *generalization*) associated with respective coding concepts (i.e., *sequences*, *loops*, *parallelism*, *events*, *conditionals*, *operators*, and *data/variables*), and contextualized in the STEM disciplines. The overlapping practical co-existence of coding concepts and CT processes in different STEM disciplines (see Study 1, Figure 2) formed the basis of the curriculum's content design approach.

CT processes were approached as cognitive counterparts of the more practical coding concepts. To offer a low-threshold first introduction to these concepts, they should initially be introduced in an unplugged modality without the use of any digital device. Later in, they should be trained in a plugged-in modality, using age-appropriate educational programming software. In addition, embodied and game-based learning methods were integrated into the didactical approaches of the curriculum's activities to increase motivation and active learning. The proposed curriculum was designed for 3rd and 4th graders and aimed at providing students with a broader perspective on the applicability of CT and coding in real-life, as a problem-solving technique.

In this study, the conceptualization and rough outline of the curriculum were demonstrated. The unplugged units of the curriculum were presented and pilot-evaluated in Studies 2 and 3.

The elaborated curriculum outline was presented in Study 4, along with its pilot and efficacy evaluations, presented in Studies 4 and 6, respectively.

8.2 Findings of Study 2: Initial evaluation of the unplugged CT games with adults

In Study 2, the unplugged units of the curriculum were presented in detail. The inspiration for the three games of the life-size board games series *Crabs & Turtles: A Series of Computational Adventures* (namely, *The Treasure Hunt*, *The Race*, and *Patterns*) was described, along with the components of the games and their instructions for play. Additionally, the results of a 2-phase empirical evaluation procedure were reported.

The aim of the games is to foster CT in elementary school children. To evaluate the games, an iterative user-centred development process was followed. The purpose of this initial evaluation was to explore the feasibility and user experience during play. The evaluation was conducted with two different adult samples, a sample of university students (*n*=17) and a sample of gamification experts and educators (*n*=19). The selection of adult samples for this evaluation stage aimed at exploring possible required adjustments to the games before evaluating them with the target population of elementary school children.

At phase 1, the overall game experience was examined after playing all three games, while at phase 2, game-specific experience was investigated after each of the three games. Quantitative feedback gathered by established game experience questionnaires (GEQ; Poels et al., 2007) and qualitative feedback provided in written and verbal formats were incorporated into the games' next version. The quantitative analysis results revealed an overall positive perception of the games, which were primarily perceived as a playful activity and to a lesser degree as a learning activity. The qualitative feedback led to minor adjustments that resulted in an improved version of the games.

After evaluating the games with adult samples, an evaluation with the actual target population of elementary school children was conducted, presented in Study 3.

8.3 Findings of Study 3: Evaluation of the unplugged CT games with students

In Study 3, the evaluation of users' gaming experience was further investigated with a sample of 70 elementary school children. After integrating the valuable feedback provided by gamification experts and educators during Study 2, the games' final version was designed and underwent a quantitative evaluation with the actual target group. The three games that constitute *Crabs & Turtles: A Series of Computational Adventures* were separately evaluated as individual game entities by utilizing the same game experience questionnaire (GEQ; Poels et al., 2007) used in Study 2.

The results of Study 3 revealed an overall positive perception of the games. Children reported feeling competent and immersed during the three games and experienced positive affect, while tension and negative affect ratings were low. The games were primarily perceived as a playful activity, and the design elements' quality was rated high. These results substantiated the games' appropriateness as playful activities for the introduction of CT concepts to elementary school children.

Nevertheless, the challenge during play was rated low for each one of the three games. For that reason, the games' instructions were enriched with a set of alternative game-play instructions that allow adaptations based on the number of children participating in the game and the available time for play. These adaptations facilitate the selection of difficulty levels and therefore allow for a more challenging players' experience.

The overall positive evaluation of the games in Study 3 succeeded to replicate our results from Study 2 at the target population and allowed to move forward to the next evaluation phase of the games, as part of a CT course intervention, investigating their cognitive and educational value when teaching CT in elementary school. This pilot evaluation phase was presented in Study 4.

8.4 Findings of Study 4: Pilot evaluation of the CT curriculum – Investigation of CT cognitive correlates

In Study 4, the pilot evaluation of the proposed in Study 1 CT curriculum was conducted. After the iterative development of the *Crabs & Turtles* games and the games' experience evaluation with different age groups, they were integrated into the newly developed curriculum, called *Verstehen wie Computer denken* [Understanding how computers work]. The curriculum consisted of ten 90-minutes lessons separated into four distinct modules (see also Study 4, Figure 1, and Study 6, Figure 2).

The course was evaluated in 4 different academies of the HCAP (Rothenbusch et al., 2016) with 31 3rd and 4th graders. The evaluation followed a pre-/post-test design, using standardized cognitive assessments for various cognitive abilities. Additionally, a CT assessment tool was used, which was derived from the adaptation of an already existing and validated for an older age group CT assessment (Román-González, Pérez-González, et al., 2017). Moreover, the cognitive correlates of CT were investigated by observing associations between students' performance at the cognitive assessments and the CT assessment.

Results indicated the effectiveness of the course with a significant students' CT performance increase from pre- to post-test. Verbal and non-verbal visuospatial reasoning skills also increased from pre- to post-test. These results replicate results from a study on an older sample of middle and secondary school children (Román-González, Pérez-González, et al., 2017) that showed associations of CT with spatial, reasoning, and problem-solving abilities. Additionally, significant associations between CT and other specific cognitive abilities were observed. CT performance was associated with complex arithmetic abilities. These results do not replicate the results of a previous study on a sample of middle and secondary school children (Román-González, Pérez-González, et al., 2017).

The findings of this study, compared to the findings on older children (Román-González, Pérez-González, et al., 2017), revealed that CT associations with other cognitive abilities are not consistent across different age groups. Non-verbal visuospatial abilities seem consistently correlated to CT from elementary to high-school level. However, numerical abilities seem more related to CT in elementary school level than in middle-school or high-school (Román-González, Pérez-González, et al., 2017). Similarly to numerical abilities, results regarding the associations of CT with verbal abilities did not replicate the results of a similar study with older children (Román-González, Pérez-González, et al., 2017). Moreover, in this study, no correlations of CT with verbal reasoning ability was observed. In Study 5 of this dissertation, however, where the same correlational analysis performed in Study 4 was replicated on a larger sample of 197 children, verbal reasoning abilities are significantly correlated to CT. This inconsistency between study results occurred most probably because of the smaller sample size of Study 4 and can therefore be interpreted as a statical power issue.

The *abbreviated CTt* used in this study to measure CT in elementary school children was an adaptation of a validated CT assessment tool for older students (*CTt*; Román-González, Pérez-González, et al., 2017). The lack of CT assessment tools for the target age group of elementary children during the time this study was conducted led to this adaptation. The use of the *abbreviated CTt* seemed as regards duration of completion and items' difficulty feasible to use with elementary school students, without any ceiling or floor effects. For that reason, the *abbreviated CTt* was further used in Studies 5 and 6. In Study 5, the correlational analysis presented in Study 4 was replicated with a larger sample size utilizing the same assessment tools and therefore, the *abbreviated CTt*. In Study 6, the pilot study (Study 4) was followed by a randomized field trial with a control group, with a larger sample size, following the same pre-/post-test protocol, aiming to investigate the effectiveness of the CT curriculum further.

8.5 Findings of Study 5: Investigation of CT cognitive correlates

In Study 5, the correlational study investigating associations of CT with other cognitive abilities as conducted previously (Study 4) was replicated with a larger sample of 197 elementary school children. Though there have been several studies investigating the cognitive correlates of programming abilities in different age groups (Jones & Burnett, 2008; Prat et al., 2020), correlational studies investigating CT as a cognitive ability in association with other cognitive abilities are limited. There have been studies investigating cognitive associations of CT in pre-school (e.g., Marinus et al., 2018), elementary (Città et al., 2019; associations of CT with mental rotation abilities), middle and high-school (e.g., Román-González, Pérez-González, et al., 2017) students, as well as university students and adults (Ambrosio et al., 2014). However, to the best of my knowledge, there have been no studies reporting associations of CT with a range of other cognitive abilities in elementary school level. The aim of this study was to define the construct of CT at this age group cognitively and therefore contribute to its definition and assessment. Along the correlational analysis, the study provided initial positive results for the reliability of a CT assessment tool for elementary school students.

To assess CT, we used the *abbreviated CTt* utilized in Study 4, along with all the other cognitive assessment tools of this study. The adaptation of the original *CTt* (Román-González, Pérez-González, et al., 2017) revealed acceptable reliability in Study 4, results which were also replicated in Study 5. The *abbreviated CTt*, consisting of 21 items and administered as a 20-minutes speeded test, showed comparable psychometric attributes as the original *CTt,* suggesting that it can be used as a reliable CT assessment on elementary school students.

The correlational analysis in Study 5 revealed differences in the cognitive interdependencies of CT across different age groups. The results indicated significant associations between CT and numerical, verbal, and non-verbal visuospatial abilities (for a visualization, see Study 5, Figure 4); however, only 24% of the variance of *CTt* performance was explained by the performance on all the other cognitive tests. This further substantiates the notion of the author's approach to cognitively define CT as a specific cognitive ability that builds upon several other cognitive abilities, which have not yet been extensively investigated in correlation to CT.

Results of Study 5 revealed: i. a weak association of CT with numerical abilities, ii. a moderate association with verbal reasoning abilities, and iii. a weak association with non-verbal visuospatial abilities. The first association replicated the results of the pilot study (Study 4) on the association of CT with simple and more complex numerical abilities, although similar studies on an older sample of middle and high-school students (Román-González, Pérez-González, et al., 2017) did not provide similar evidence. The results of Study 5 are in line with the initial results of the pilot study (Study 4), substantiating the argument that numerical abilities are a prerequisite for thinking computationally at an early stage of cognitive development.

The second association that of CT with verbal abilities did not replicate the results of the pilot study (Study 4) where no significant association between CT and verbal reasoning was revealed. Nevertheless, this positive association replicates results of previous studies conducted on samples of older and younger age groups investigating cognitive associations of CT with verbal abilities (e.g., Good & Howland, 2017; Marinus et al., 2018; Prat et al., 2020; Román-González, Pérez-González, et al., 2017). This, however, is not the case with young adults, where verbal ability seems not to be associated with programming experience and CT (Helmlinger et al., 2020). These results indicate that language is an essential factor in the development of CT at a younger age, but not necessarily after secondary education.

The third association revealed by the correlational analysis is the one of CT with non-verbal visuospatial abilities. This weak association substantiates the results of the pilot study (Study 4) and results of several other studies conducted on similar (Città et al., 2019) or older age groups (e.g., Ambrósio et al., 2015; Jones & Burnett, 2008; Parkinson & Cutts, 2018; Román-González, Pérez-González, et al., 2017; M. Werner, 2020). These results indicate that non-verbal visuospatial reasoning abilities are consistently and across educational levels a predictor of thinking computationally.

The results of this study regarding the cognitive associates of CT in elementary school level complement the nomological network of the cognitive abilities related to the development of CT across age groups and therefore support a cognitive definition of CT as a psychological construct (for a visualization, see Study 5, Figure 5).

8.6 Findings of Study 6: Effectiveness evaluation of the CT curriculum

In Study 6, the pilot evaluation of the CT curriculum (Study 4) was replicated with a sample of 197 elementary school children, in a randomized field trial with a control group, following the same pre-/post-test protocol. This last phase of the evaluation aimed at investigating the effectiveness of the CT curriculum (for an elaborated description of the curriculum content, see Study 6, Figure 2). To evaluate the curriculum in a real classroom setting, the training in this phase was delivered by trained instructors and not the developers of the course as in the pilot phase of the evaluation (Study 4). The effectiveness of the curriculum was measured utilizing the various cognitive and CT assessments used in Study 4.

Evaluation of the pre- and post-test differences revealed a larger improvement in students' CT performance for the intervention group compared to the control group. This significant positive effect of the CT curriculum on students' CT abilities replicates the findings of the pilot study (Study 4) where no control group design was implemented. These results substantiate the previous evidence on the effectiveness of the CT curriculum even when delivered by instructors in real classroom settings.

The small effect size of our results though not desirable is in line with existing literature on the usually small effect sizes reported for CT interventions with block-programming languages (Chiu & Tsuei, 2020). Since similar empirical studies on the effectiveness of CT interventions for elementary school are limited (Upadhyaya et al., 2020), this study could contribute to the body of research focusing on the design and evaluation of effective educational materials for fostering and developing CT.

9 Conclusions

In this section, the summary of results presented in section 8 are recapped, structured in two subsections based on this thesis's objectives, namely: i. the *Curriculum design and development for fostering CT* (see section 5.1), and ii. the *Cognitive correlates of CT and its assessment* (see section 5.2). First, the CT curriculum design approach, the development procedure, and the evaluation results will be discussed (section 9.1). Second, the investigation of the cognitive correlates of CT, along with a CT assessment tool for the elementary school level, will be recapitulated (section 9.2).

9.1 Curriculum for fostering CT

The design and evaluation of the CT curriculum developed in this thesis was based on a literature review on the most recent didactical approaches suggested for fostering CT to elementary school children. Since CT lacks a concrete and widely accepted definition, as a working definition for the six enclosed studies, the definition of Shute et al. (2017) was considered. This definition interprets CT as the underpinning construct of effective and efficient problem-solving, applied in contexts within and beyond CS. The design of the CT curriculum required the set of concrete components to be fostered, a transparent context of introduction and application of the components, and didactical methods relevant for introducing complex concepts to elementary school students (Study 1).

There have been specific CT processes identified in the literature which are cognitively supporting more practical skills relevant to coding. Though these CT processes can be well-expressed within the context of CS and coding, they are not limited to this context. However, for the purposes of this course, coding served well as an educational context for introducing and fostering CT. To provide a broader perspective on the wide applicability of CT, the activities incorporated in the curriculum focused on different STEM topics. Additionally, to motivate students' participation and offer a low-threshold introduction to complex concepts, game-based learning and embodiment were facilitated with the development of the life-size board games *Crabs & Turtles*.

The games went through an iterative development procedure, having been evaluated in different phases and with different samples of players for the game experience they offer (Studies 2 and 3). After integrating feedback from the different evaluation phases, the games were integrated into the proposed CT curriculum and underwent a pilot and an effectiveness evaluation (Studies 4 and 6 respectively) with more than 200 elementary school students.

Results of the different evaluation studies indicated that the design approach of the CT curriculum had a positive effect on students' CT performance. The blended approach of a low-threshold embodied, and unplugged introduction of CT and coding concepts with playful activities and the plugged-in transfer of the same concepts in visual block-programming environments of different output modalities (i.e., *Scratch*: small-scale games and interactive applications, *S4A*: haptic and visual applications with a microcontroller, *Open Roberta Lab:* robot simulations) proved itself efficient. Given the fact that the evaluations of the CT curriculum focused on a sample of students attending an extracurricular enrichment program, future research is needed to evaluate whether similar results would occur in formal classroom settings.

Taken together, the findings of the evaluation studies support the idea of a blended unplugged and plugged-in approach on teaching concrete CT concepts using coding activities as a vehicle for applying them in various STEM domains. The evaluated curriculum is a Hector Core Course[24] at the Hector Children's Academy Program[25] (Rothenbusch et al., 2016) for talented children and has been offered as an extracurricular course, across the 66 academies in Baden-Württemberg, since 2018.

Additionally, the proposed educational life-size board games developed and evaluated for the scope of this curriculum seem to facilitate a playful introduction to CT and coding concepts. Their publication as an OER allows for open access and further use of the materials in various educational contexts, assisting a direct transfer of research development to educational practice. Furthermore, the multi-phase evaluation studies of the developed curricula have gone some way towards enhancing the iterative development of CT curricula, evaluated for their effectiveness in real-class conditions, and providing robust statistical information that will allow future comparative studies with the respective age group.

[24] Hector Core Course are course developed for the Hector Children's Academy Program and are offered by trained instructors across the 66 academies of the program, after being evaluated for their effectiveness.
[25] https://hector-kinderakademie.de

9.2 Cognitive correlates and assessment of CT

The cognitive definition of CT that this thesis seeks to enhance was pursued by focusing on the development of the construct at the population of elementary school children, and derived from evaluating associations of performance on a CT test and performance on other tests of well established cognitive constructs. Since there was no validated CT test for the age group of elementary students at the period, this PhD research took place, a CT test for older students (*CTt*; Román-González, Pérez-González, et al., 2017) was adapted and used for the purposes of this research. The *abbreviated CTt* showed adequate reliability in both Studies 4 and 5, as well as similar psychometric properties to the original *CTt*. The assessment will be openly accessible via the OSF[26], and an implication of this is hopefully the use of the proposed *abbreviated CTt* for future research on elementary school students, where no validated CT assessment tools are yet widely available.

The results of the association analyses between CT and other cognitive abilities (i.e., numerical abilities, verbal reasoning, and non-verbal reasoning abilities) complements the nomological network of CT, by providing empirical evidence on the cognitive underpinnings of CT within the underinvestigated population of elementary school students (for a visualization, see Study 5, Figure 5). Though the correlational analyses revealed some different cognitive associations of CT at the elementary school level compared to those observed in older age groups, they also substantiated others.

Non-verbal visuospatial abilities seem to be consistently correlated to CT from elementary to high-school level (Román-González, Pérez-González, et al., 2017). Numerical abilities, however, seem more relevant to CT in elementary than in middle-school or high-school level (Román-González, Pérez-González, et al., 2017). This differentiation of numerical correlations with CT across age could be explained by the fact that at an early age, numerical abilities are prerequisites for CT, while later on when a certain threshold of numerical ability is achieved through formal education, numerical abilities are not decisive for the development of CT (see Helmlinger et al., 2020; Prat et al., 2020 for comparable results in adult populations). Verbal abilities seems to be associated with CT in elementary school children as well as in younger and older populations (Good & Howland, 2017; Marinus et al., 2018; Prat et al., 2020; Román-González, Pérez-González, et al., 2017). Nevertheless, in young adults, verbal ability seems not

[26] The OSF link provided in Study 5 will be openly accessible after the manuscript's publication.

to be associated with programming experience and CT any more (Helmlinger et al., 2020). The decreasing association of CT with verbal abilities in older samples might be explained by the fact that in younger age, the use of language serves as a scaffolding cognitive strategy to read and formulate algorithms while this strategy becomes less relevant when getting older and thus more experienced. The observed associations indicate that language is an important factor for CT development at a younger age, but not necessarily after secondary education.

In summary, at the elementary school level, CT, as measured by the *abbreviated CTt*, is moderately to weakly associated with numerical, verbal, and non-verbal visuospatial reasoning abilities. However, variance in performance on the *abbreviated CTt* was only partially explained by these other cognitive constructs. This provides further evidence for the argument that CT is a specific cognitive ability that builds on and recruits a convolute of several other cognitive abilities, which are not yet investigated and understood comprehensively in relation to CT. Nevertheless, the results of the current study clearly indicate that CT performance is more than just the sum of the assessed other cognitive constructs.

As such, and with the educational value it has been assigned, CT should be further investigated as a unique cognitive ability, taking into consideration its associations with other cognitive and non-cognitive factors until it is well defined and reliably measured across educational levels. Such factors proposed already are personality and self-efficacy (Román-González et al., 2018b), executive functions (Robertson et al., 2020), and creative thinking (Scherer et al., 2018).

10 Future Perspectives

The work related to CT, and more specifically, the cognitive investigation, the curricula design and effectiveness as well as the assessment of CT are still in their infancy, despite the considerable amount of research being conducted worldwide (Lockwood & Mooney, 2017; Román-González et al., 2018). CT as a cognitive construct needs further investigation that will allow for more appropriate assessment tools design and therefore, more effective educational interventions for fostering CT. The presented CT curriculum design approach and the multi-phase evaluation procedure for measuring the intervention's effectiveness outline design and evaluation perspectives on future CT curricula addressed to young children at elementary school level. The proposed CT assessment tool and the investigated cognitive correlates of CT provide a promising tool for assessing CT in elementary school children, complement the research on the cognitive embedding of CT at the respective age group and provide a robust basis for further research on the definition of CT as a unique cognitive ability.

To further advance research on CT, future similar studies are required in order to elucidate the development of CT in younger and older age groups (like pre-schoolers and university students) where the research on the cognition of CT is comparatively limited as in elementary school (Upadhyaya et al., 2020). That would allow for a more comprehensive picture of the cognitive development of CT across age-groups. In this direction, a replication study (Marinus et al., 2018) on the cognitive correlates and assessment of CT with 5 and 6 years old students was conducted in 2019 and awaits analysis. Similarly, a study on the cognitive correlates of CT with numerical/mathematical abilities and their differential prediction of coding performance has been conducted in university students, and part of the results have been presented by Werner (2020). Besides, future studies are required to substantiate the argument that CT is a unique cognitive ability, that even though it recruits and seems to rest on a convolute of other cognitive abilities, may also associate with factors that are not yet extensively investigated in relation to CT, like creative thinking, executive functions, and non-cognitive behavioural factors.

Similarly, CT assessment research has a long way to go until reliable CT assessment tools for different age-groups are developed. Further cross-validations studies of different CT assessment tools are needed in order to ensure the tools' psychometric quality. A first step in this regard is a study conducted earlier this year, translating in German and validating the *BCTt*

(Zapata-Cáceres et al., 2020), a newly developed CT assessment for primary school children that had not been yet developed at the time this study took place. Future studies are planned for cross-validating *BCTt* with the *abbreviated CTt.*

References

Aggarwal, A., Gardner-Mccune, C., & Touretzky, D. S. (2017). Evaluating the effect of using physical manipulatives to foster computational thinking in elementary school. *Proceedings of the Conference on Integrating Technology into Computer Science Education, ITiCSE*, 9–14. https://doi.org/10.1145/3017680.3017791

Ambrosio, A. P., da Silva Almeida, L., Macedo, J., & Franco, A. (2014). Exploring Core Cognitive Skills of Computational Thinking. *Psychology of Programming Interest Group Annual Conference 2014 Proceedings*, *July*, 25–24. http://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf%0Ahttp://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.698.1911&rep=rep1&type=pdf

Ambrósio, A. P., Xavier, C., & Georges, F. (2015). Digital ink for cognitive assessment of computational thinking. *Proceedings - Frontiers in Education Conference, FIE*, *February*. https://doi.org/10.1109/FIE.2014.7044237

Angeli, C., Voogt, J., Fluck, A., Webb, M., Cox, M., Malyn-Smith, J., & Zagami, J. (2016). International Forum of Educational Technology & Society A K-6 Computational Thinking Curriculum Framework: Implications for Teacher Knowledge. *Journal of Educational Technology & Society*, *19*(3), 47–57.

Armoni, M. (2016). COMPUTING IN SCHOOLS Computer science, computational thinking, programming, coding. *ACM Inroads*, *7*(4), 24–27. https://doi.org/10.1145/3011071

Astrachan, O., & Briggs, A. (2012). *The CS Principles Project*. *3*(2), 38–42.

Barr, D., Harrison, J., & Conery, L. (2011). Computational Thinking: A Digital Age Skill for Everyone. *Learning and Leading with Technology*, *38*(6), 20–23.

Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12. *ACM Inroads*, *2*(1), 48–54. https://doi.org/10.1145/1929887.1929905

Bergin, S., & Reilly, R. (2006). Predicting introductory programming performance: A multi-institutional multivariate study. *Computer Science Education*, *16*(4), 303–323. https://doi.org/10.1080/08993400600997096

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

Brackmann, C. P., Román-González, M., Robles, G., Moreno-León, J., Casali, A., & Barone, D. (2017). *Development of Computational Thinking Skills through Unplugged Activities in Primary School*. 65–72. https://doi.org/10.1145/3137065.3137069

Brennan, K., & Resnick, M. (2012a). New frameworks for studying and assessing the development of computational thinking. *Annual American Educational Research Association Meeting, Vancouver, BC, Canada*, 1–25. http://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf

Brennan, K., & Resnick, M. (2012b). Using artifact-based interviews to study the development of computational thinking in interactive media design. *Proceedings of the Annual American Educational Research Association Meeting (AERA)*. http://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf

Brown, N. C. C., Sentance, S. U. E., Crick, T. O. M., & Humphreys, S. (2014). Restart: The Resurgence of Computer Science in UK Schools. *ACM Transactions on Computing Education*, *14*(2), 1–22. https://doi.org/10.1145/2602484

Byrne, P., & Lyons, G. (2001). The effect of student attributes on success in programming. *ACM SIGCSE Bulletin*, *33*(3), 49–52. https://doi.org/10.1145/507758.377467

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. Cambridge University Press. https://doi.org/10.1177/001698629904300207

Chen, G., Shen, J., Barth-Cohen, L., Jiang, S., Huang, X., & Eltoukhy, M. (2017). Assessing elementary students' computational thinking in everyday reasoning and robotics programming. *Computers & Education*, *109*, 162–175. https://doi.org/10.1016/j.compedu.2017.03.001

Chiprianov, V., & Gallon, L. (2016). Introducing Computational Thinking to K-5 in a French Context. *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '16*, 112–117. https://doi.org/10.1145/2899415.2899439

Chiu, J.-I., & Tsuei, M. (2020). *Meta-Analysis of Children's Learning Outcomes in Block-Based Programming Courses* (pp. 259–266). https://doi.org/10.1007/978-3-030-60703-6_33

Città, G., Gentile, M., Allegra, M., Arrigo, M., Conti, D., Ottaviano, S., Reale, F., & Sciortino, M. (2019). The effects of mental rotation on computational thinking. *Computers & Education*, *141*(July), 103613. https://doi.org/10.1016/j.compedu.2019.103613

Curzon, P., Dorling, M., Selby, C., Woollard, J., & Ng, T. (2014). *Developing computational thinking in the classroom: a framework*. *June*. http://eprints.soton.ac.uk/369594/10/DevelopingComputationalThinkingInTheClassroomaFramework.pdf

Dierbach, C., Hochheiser, H., Collins, S., Jerome, G., Ariza, C., Kelleher, T., Kleinsasser, W., Dehlinger, J., & Kaza, S. (2011). A Model for Piloting Pathways for Computational Thinking in a General Education Curriculum. *Development*, *15*(5), 257–262. https://doi.org/10.1145/1953163.1953243

Falkner, K., Vivian, R., & Falkner, N. (2014). The Australian digital technologies curriculum: Challenge and opportunity. *Conferences in Research and Practice in Information Technology Series*, *148*(January), 3–12.

Garcia-Peñalvo, F. J. (2016). What Computational Thinking Is. *Journal of Information Technology Research*, *9(3)*, *v–vi*(October).

Good, J., & Howland, K. (2017). Programming language, natural language? Supporting the diverse computational activities of novice programmers. *Journal of Visual Languages and Computing*, *39*(October 2016), 78–92. https://doi.org/10.1016/j.jvlc.2016.10.008

Grover, S., & Pea, R. (2013). Computational Thinking in K-12: A Review of the State of the Field. *Educational Researcher*, *42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

Helmlinger, B., Sommer, M., Feldhammer-Kahr, M., Wood, G., Arendasy, M. E., & Kober, S. E. (2020). Programming experience associated with neural efficiency during figural reasoning. *Scientific Reports*, *10*(1), 1–14. https://doi.org/10.1038/s41598-020-70360-z

Howland, K., & Good, J. (2015). Learning to communicate computationally with Flip: A bi-modal programming language for game creation. *Computers and Education*, *80*, 224–240. https://doi.org/10.1016/j.compedu.2014.08.014

Jones, S., & Burnett, G. (2008). Spatial Ability and Learning to Program. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, *4*(1), 47–61. https://doi.org/10.17011/ht/urn.200804151352

Jovanov, M., Stankov, E., Mihova, M., Ristov, S., & Gusev, M. (2016). Computing as a new compulsory subject in the Macedonian primary schools curriculum. *IEEE Global Engineering Education Conference, EDUCON*, *10-13-Apri*(April), 680–685. https://doi.org/10.1109/EDUCON.2016.7474623

Kalelioğlu, F., Gülbahar, Y., & Kukul, V. (2016). A Framework for Computational Thinking Based on a Systematic Research Review. *Baltic J . Modern Computing*, *4*(3), 583–596.

Koh, K. H., Nickerson, H., Basawapatna, A., & Repenning, A. (2014). Early validation of computational thinking pattern analysis. *Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education - ITiCSE '14*, 213–218. https://doi.org/10.1145/2591708.2591724

Kong, S.-C., & Abelson, H. (2019). Computational Thinking Education. In S.-C. Kong & H. Abelson (Eds.), *Computational Thinking Education*. Springer Singapore. https://doi.org/10.1007/978-981-13-6528-7

Leifheit, L., Tsarava, K., Ninaus, M., & Moeller, K. (2018). *"Verstehen wie Computer denken" - Ein Trainingsprogramm zur Förderung von informatischem Denken und systematischen Problemlösefähigkeiten besonders begabter Kinder im Grundschulalter*. Reihe Hector Core Courses.

Lockwood, J., & Mooney, A. (2017). *Computational Thinking in Education : Where does it Fit ? A systematic literary review A systematic literary review*. *March*, 1–58.

Marinus, E., Powell, Z., Thornton, R., McArthur, G., & Crain, S. (2018). Unravelling the Cognition of Coding in 3-to-6-year Olds. *Proceedings of the 2018 ACM Conference on International Computing Education Research - ICER '18*, *August*, 133–141. https://doi.org/10.1145/3230977.3230984

McCoy, L. P., & Burton, J. K. (1988). The relationship of computer programming and mathematics in secondary students. *Computers in the Schools*, *4*(3–4), 159–166.

McGill, M. M., & Decker, A. (2020). A Gap Analysis of Statistical Data Reporting in K-12 Computing Education Research. *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 591–597. https://doi.org/10.1145/3328778.3366842

Moreno-León, J., Robles, G., & Román-González, M. (2015). Dr. Scratch: Automatic Analysis of Scratch Projects to Assess and Foster Computational Thinking. *RED. Revista de Educación a Distancia*, *15*(46), 1–23. https://doi.org/10.6018/red/46/10

Moreno-Leon, J., Roman-Gonzalez, M., & Robles, G. (2018). On computational thinking as a universal skill: A review of the latest research on this ability. *2018 IEEE Global Engineering Education Conference (EDUCON)*, 1684–1689. https://doi.org/10.1109/EDUCON.2018.8363437

Mühling, A., Ruf, A., & Hubwieser, P. (2015). Design and First Results of a Psychometric Test for Measuring Basic Programming Abilities. *Proceedings of the Workshop in Primary and Secondary Computing Education*, 2–10. https://doi.org/10.1145/2818314.2818320

Nowaczyk, R. H. (1983). *Cognitive Skills Needed in Computer Programming*. https://www.learntechlib.org/p/136288

OECD. (1970). Pisa 2021 Mathematics Framework(Draft). *Journal of Chemical Information and*

Modeling, *53*(9), 1689–1699. https://doi.org/10.1017/CBO9781107415324.004

Papert, S. (1999). Introduction: What is Logo? And Who Needs It? In *Logo Philosophy and Implementation*. LCSI. http://www.microworlds.com/support/logo-philosophy-papert.html

Papert, S., & Harel, I. (1991). *Constructionism*. Ablex Publishing Corporation.

Parkinson, J., & Cutts, Q. (2018). *Investigating the Relationship Between Spatial Skills and Computer Science*. 106–114. https://doi.org/10.1145/3230977.3230990

Parkinson, J., & Cutts, Q. (2019). Chairs' AWARD. *ACM Inroads*, *10*(1), 64–73. https://doi.org/10.1145/3306151

Pea, R. D., & Kurland, D. M. (1984). On the cognitive effects of learning computer programming. *New Ideas in Psychology*, *2*(2), 137–168. https://doi.org/10.1016/0732-118X(84)90018-7

Perković, L., Settle, A., Hwang, S., & Jones, J. (2010). A framework for computational thinking across the curriculum. *Proceedings of the Fifteenth Annual Conference on Innovation and Technology in Computer Science Education - ITiCSE '10*, 123. https://doi.org/10.1145/1822090.1822126

Poels, K., de Kort, Y., & Ijsselsteijn, W. (2007). *FUGA - The fun of gaming: Measuring the human experience of media enjoyment. Deliverable D3.3: Game Experience Questionnaire*.

Prat, C. S., Madhyastha, T. M., Mottarella, M. J., & Kuo, C. H. (2020). Relating Natural Language Aptitude to Individual Differences in Learning Programming Languages. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-60661-8

Qualls, J. A., & Sherrell, L. B. (2010). Why computational thinking should be integrated into the curriculum. *Journal of Computing Sciences in Colleges*, *25*(5), 66–71.

Román-González, M., Moreno-León, J., & Robles, G. (2017). Complementary Tools for Computational Thinking Assessment. *International Conference on Computational Thinking Education 2017*, *July*.

Román-González, M., Moreno-León, J., & Robles, G. (2019). Combining Assessment Tools for a Comprehensive Evaluation of Computational Thinking Interventions. In *Computational Thinking Education* (pp. 79–98). Springer Singapore. https://doi.org/10.1007/978-981-13-6528-7_6

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

Román-González, M., Pérez-González, J. C., Moreno-León, J., & Robles, G. (2018). 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

Rothenbusch, S., Zettler, I., Voss, T., Losch, T., & Trautwein, U. (2016). Exploring reference group effects on teachers' nominations of gifted students. *Journal of Educational Psychology*, *108*(6), 883–897. https://doi.org/10.1037/edu0000085

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

Seiter, L., & Foreman, B. (2013). Modeling the learning progressions of computational thinking of primary grade students. *Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research - ICER '13*, 59.

https://doi.org/10.1145/2493394.2493403

Sentance, S., & Csizmadia, A. (2017). Computing in the curriculum: Challenges and strategies from a teacher's perspective. *Education and Information Technologies*, *22*(2), 469–495. https://doi.org/10.1007/s10639-016-9482-0

Settle, A., Franke, B., Hansen, R., Spaltro, F., Jurisson, C., Rennert-May, C., & Wildeman, B. (2012). *Infusing computational thinking into the middle- and high-school curriculum*. 22. https://doi.org/10.1145/2325296.2325306

Settle, A., Goldberg, D. S., & Barr, V. (2013). *Beyond computer science*. *July*, 311. https://doi.org/10.1145/2462476.2462511

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

Sneider, C., Stephenson, C., Schafer, B., & Flick, L. (2014). Computational Thinking in High School Science Classrooms: Exploring the Science "Framework" and "NGSS." *Science Teacher*, *81*(5), 53–59. https://www.learntechlib.org/p/155904

Sullivan, A., Kazakoff, E. R., & Bers, M. U. (2013). The wheels on the bot go round and round: Robotics curriculum in pre-kindergarten. *Journal of Information Technology Education*, *12*, 203–219. http://www.jite.org/documents/Vol12/JITEv12IIPp203-219Sullivan1257.pdf

Tang, K. (2019). A Content Analysis of Computational Thinking Research : An International Publication Trends and Research Typology. *The Asia-Pacific Education Researcher*. https://doi.org/10.1007/s40299-019-00442-8

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

Tran, Y. (2019). Computational Thinking Equity in Elementary Classrooms: What Third-Grade Students Know and Can Do. *Journal of Educational Computing Research*, *57*(1), 3–31. https://doi.org/10.1177/0735633117743918

Tsarava, K., Leifheit, L., Ninaus, M., Román-González, M., Butz, M. V., Golle, J., Trautwein, U., & Moeller, K. (2019). Cognitive Correlates of Computational Thinking. *Proceedings of the 14th Workshop in Primary and Secondary Computing Education on - WiPSCE'19*, *October*, 1–9. https://doi.org/10.1145/3361721.3361729

Upadhyaya, B., McGill, M. M., & Decker, A. (2020). A longitudinal analysis of k-12 computing education research in the united states: Implications and recommendations for change. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, 605–611. https://doi.org/10.1145/3328778.3366809

Van Dyne, M., & Braun, J. (2014). Effectiveness of a computational thinking (CS0) course on student analytical skills. *Proceedings of the 45th ACM Technical Symposium on Computer Science Education - SIGCSE '14*, *May*, 133–138. https://doi.org/10.1145/2538862.2538956

Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2013). Challenges to learning and schooling in the digital networked world of the 21st century. *Journal of Computer Assisted Learning*, *29*(5), 403–413. https://doi.org/10.1111/jcal.12029

Voogt, Joke, Fisser, P., Good, J., Mishra, P., & Yadav, A. (2015). Computational thinking in compulsory education: Towards an agenda for research and practice. *Education and Information*

Technologies, *20*(4), 715–728. https://doi.org/10.1007/s10639-015-9412-6

Wang, P., Fessard, G., & Wang, P. (2019). *Are There Differences in Learning Gains When Programming a Tangible Object or a Simulation ? Are There Differences in Learning Gains When Programming a Tangible Object or a Simulation ? July*. https://doi.org/10.1145/3304221.3319747

Wang, P. S. (2015). *From Computing to Computational Thinking* (1st ed.). Chapman and Hall/CRC.

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

Weintrop, D., Beheshti, E., Horn, M. S., Orton, K., Trouille, L., Jona, K., & Wilensky, U. (2014). Interactive Assessment Tools for Computational Thinking in High School STEM Classrooms. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, *136 LNICST*, 22–25. https://doi.org/10.1007/978-3-319-08189-2_3

Weintrop, D., & Wilensky, U. (2015). Using commutative assessments to compare conceptual understanding in blocks-based and text-based programs. *ICER 2015 - Proceedings of the 2015 ACM Conference on International Computing Education Research*, *August*, 101–110. https://doi.org/10.1145/2787622.2787721

Werner, L., Denner, J., Campe, S., & Kawamoto, D. C. (2012). The fairy performance assessment. *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education - SIGCSE '12*, 215. https://doi.org/10.1145/2157136.2157200

Werner, M. (2020). *Computational Thinking in Beziehung zu seinen verwandten psychologischen Konstrukten*. University of Tübingen.

Wiebe, E., Mott, B. W., London, J., Boyer, K. E., Aksit, O., & Lester, J. C. (2019). Development of a lean computational thinking abilities assessment for middle grades students. *SIGCSE 2019 - Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 456–461. https://doi.org/10.1145/3287324.3287390

Wing, J. M. (2006a). Computational Thinking. *Theoretical Computer Science*, *49*(3), 33–35. https://doi.org/https://www.cs.cmu.edu/~15110-s13/Wing06-ct.pdf

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

Wing, J. M. (2010). Computational Thinking: What and Why? *Thelink - The Magaizne of the Varnegie Mellon University School of Computer Science*, *March 2006*, 1–6. http://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational Thinking in Elementary and Secondary Teacher Education. *ACM Transactions on Computing Education*, *14*(1), 1–16. https://doi.org/10.1145/2576872

Yadav, A., Zhou, N., Mayfield, C., Hambrusch, S., & Korb, J. T. (2011). Introducing computational thinking in education courses. *Proceedings of the 42nd ACM Technical Symposium on Computer Science Education - SIGCSE '11*, *2*, 465. https://doi.org/10.1145/1953163.1953297

Yaşar, O. (2018a). A new perspective on computational thinking. *Communications of the ACM*, *61*(7), 33–39. https://doi.org/10.1145/3214354

Yaşar, O. (2018b). Computational Thinking, Redefined. In E. Langran & J. Borup (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (Issue June, pp. 72–80). Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/primary/p/182505/

Zapata-Cáceres, M., Martín-Barroso, E., & Román-González, M. (2020). Computational Thinking Test for Beginners: Design and Content Validation. *2020 IEEE Global Engineering Education Conference (EDUCON)*, 1905–1914. https://doi.org/10.1109/EDUCON45650.2020.9125368

Acknowledgements

Throughout my PhD studies that resulted in this dissertation, I have received a great deal of professional and personal support and assistance by many human and animal beings (mainly dogs). I will here attempt to name them all properly.

I would first like to thank my supervisor, Prof. Dr. Korbinian Moeller, for accepting me as a member of his former group (the great Neurocognitive plasticity group), for providing me with his constant support, his always detailed and immediate feedback, and his endless patience! Furthermore, I wish to thank my supervisor Prof. Dr. Martin V. Butz, for always offering his valuable perspective on cognition throughout the different steps of my research studies. Just as importantly, I wish to thank Dr. Manuel Ninaus for being the most outstanding “unofficial” supervisor that one could wish for, being always a great source of support, inspiration and motivation.

Moreover, I would like to thank Prof. Dr. Marcos Román-González for always promptly providing access to his research materials and for his constant feedback on my work on computational thinking. I also wish to thank Prof. Dr. Hartmut Leuthold for providing the second assessment of this thesis on a very short notice from my side, just right before Christmas vacation and during weird corona times.

I am grateful to all my colleagues at *Ebene 5 ¾* at the Leibniz Institute für Wissensmedien (IWM) for their support while drinking coffee or *bubbly*, eating cake, and crafting! It was a great pleasure working with you all. A special thank goes to my wonderful ex office-mate Julia Wunsch, and to the lovely and supportive Silke Wortha, Roberta Barrocas, Dr. Steffi Jung, Dr.^{2} Elise Klein, and Anna -the kid- for making the office (or zoom meeting rooms) always a fun place to be. Additionally, I would like to thank the team *Wissenschaftliche Begleitung der Hector Kinderakademien* in the Hector Institute (HIB) for supporting my research all along the way. A special thanks to Kristin Funcke, and Prof. Dr. Jessika Golle for their guidance and support until the very last days of my studies, as well as to Luzia Leifheit for the time we invested collaborating. A great thanks to the *Medientechnik* of IWM and the *IT crowd* of HIB for their uncomplicated support during all my technical struggles.

My studies abroad would not have been possible without the support of my friends and colleagues Stella Patila, Dr. Christos Stergiou, Dr. Vasilios Mavroudis, and Vasilis Gkogkidis.

Thank you all! My struggle with German administrative procedures related to my PhD would not have been resolved without the help of fellow human beings like Timo and Fabian.

To my family in Tübingen, Safi and Vik, and my admirable friends Melania, Monica, Maria, Maraki, Athena & Manos, Gizem, Aggelos, Nakis, and Titika: I am forever grateful for having met you. You have all been a great source of light and happiness to me. To my *absolutely fabulous* friends from Greece, Zeta, Irene, Panagiotis, Markos, Thanos, Angela, Mpatila, Eleni and Kot: I am grateful to have always had you by my side (or a click-away), to the darkest and lightest of times, unfiltered and since time-zero.

Finally, to my family in Greece, Sonia, Thanasis, and Karlos: thank you for tolerating, unconditionally loving and supporting me. Quoting K. Gogou *„Σημασία έχει να παραμένεις άνθρωπος"* [What matters is to remain humain], and I have been learning that from you till today.

Declaration

I hereby declare that I have produced the present work independently and without the use of any aids other than those indicated. All passages that have been adopted verbatim or mutatis mutandis from published or unpublished resources are referenced as such. I also declare that I am the sole author of this PhD thesis. This work has not yet been submitted in the same, in a similar form, or in excerpts in another examination, nor has it been part of any other dissertation.

Tübingen, 16.12.2020 Katerina Tsarava