A Systematic Review of Think Alouds in Computational Thinking Research

by

Zexuan Pan

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Education

in

Measurement, Evaluation, and Data Science

Department of Educational Psychology
University of Alberta

Abstract

Computational thinking (CT) has become an essential literacy for the digital age, as it enables individuals to approach complex problems systematically and adapt to a rapidly changing technological landscape. Despite the educational benefits of CT being largely supported in the literature, the cognitive processes associated with CT are not yet well understood, restricting researchers and practitioners from designing more effective instructional strategies and personalized learning experiences. This warrants the investigation of alternative data collection methods that can provide deeper insights into individuals' cognitive processes during CT-related activities. One such method is think-aloud interviews, which involve participants verbalizing their real-time thoughts while engaging in problem-solving tasks. This systematic review examines 35 empirical studies featuring the use of think-aloud interviews in CT research. Findings show that think-aloud interviews (1) are typically conducted in Computer Science classrooms and with K-12 students; (2) are usually combined with other exploratory CT assessment tools; (3) have the potential to benefit learners with special needs and identify the competency gaps through involving diverse participants; (4) are conducted in the absence of cognitive models and standard procedures; and (5) display insufficient definitional and methodological rigor. Theoretically, this review presents a systematic assessment of the application of think-aloud interviews in CT studies and identifies gaps in existing CT-related think-aloud studies. Practically, this review serves as a reference for studying the cognitive processes during CT problem-solving and provides suggestions for CT researchers who intend to incorporate think-aloud interviews in their studies.

Preface

Our article "Insights into Computational Thinking from Think-Aloud Interviews: A Systematic Review" by Zexuan Pan, Ying Cui, Jacqueline P. Leighton, and Maria Cutumisu was published in the *Applied Cognitive Psychology* journal, volume *37*, issue 1, pages 71–95, doi: https://doi.org/10.1002/acp.4029.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Dr. Maria Cutumisu, who has been a constant source of inspiration and guidance throughout my academic journey. Her unwavering belief in my potential and persistent encouragement have significantly contributed to my intellectual development and personal growth. As a remarkable female scholar embodying confidence, perseverance, and kindness, she is a lifelong role model for me, motivating me to courageously pursue excellence in my endeavors. Moving forward in my research career, I am committed to sharing the invaluable lessons she has imparted with future colleagues and collaborators.

I would like to extend my sincere gratitude to Dr. Jacqueline Pei and Dr. Leila Ranta, members of my examining committee, as well as Dr. Mike Carbonaro, the chair of my examination, for generously contributing their time and expertise to provide insightful feedback and help me achieve this important milestone. Additionally, I would like to thank my lab mates and peers for their companionship and inspiration, which are instrumental in making my master's studies enjoyable and fulfilling. I would also like to thank Dr. Susanne Lajoie and Ms. Tianshu Li for introducing me to the field of Educational Psychology and providing me with an opportunity to explore the beauty of education research. My heartfelt appreciation is extended to my favorite musicians, Mr. Ryuichi Sakamoto and Ms. Yico Tseng, for their artistic creations that have brought me spiritual happiness and nurtured my humanistic concern for future generations.

I am grateful to my dearest mother and father for their unconditional love throughout my entire life. I also want to thank my other family members for their trust and support. I am truly blessed to have such a caring family who has been there for me every step of the way.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

AR	Augmented Reality
CHC	Cattell-Horn-Carrol
CS	Computer Science
CT	Computational Thinking
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
STEM	Science, Technology, Engineering, and Mathematics

Chapter 1 Introduction

The rapid development of computers has accelerated the digitalization of society, resulting in an increasing demand for technologically fluent workforces. Computational thinking (CT), which refers to a set of problem-solving skills and attitudes inspired from computer science (Wing, 2006), has been hailed as an important aspect of digital literacy and a fundamental ability for learners in the twenty-first century. Previous empirical studies showed that developing students' CT had positive effects on a range of cognitive and metacognitive abilities (Israel-Fishelson & Hershkovitz, 2022; Tsarava et al., 2019; Yadav et al., 2022). Integrating CT into contemporary education can also increase students' learning motivation and confidence, especially in science, technology, engineering, and mathematics (STEM, Allan et al., 2010; Kazimoglu, 2020; Sengupta et al., 2013). While the educational benefits of CT have been largely supported in the literature, the cognitive processes associated with CT remain unclear. Learners' cognitive processes that unfold during CT activities are essential and can help researchers and practitioners (a) gain a better understanding of CT in different learning contexts across developmental stages, (b) identify individual differences, and (c) provide personalized instruction that facilitates effective learning.

CT researchers usually asked questions about CT problem-solving practices, understanding, and the related cognitive processes (Tang et al., 2020). For example, (1) What are the different types of challenges that students face while working on CTSiM and what kind of supports can help them to overcome these challenges (Basu et al., 2016)? (2) How does students' CT understanding correspond to the hypothetical cognitive progression of the learning trajectories (Luo et al., 2020)? (3) What problem-solving strategies do African-American students demonstrate when solving computer science problems (Jones-Harris & Chamblee, 2017)? Although the numerical data collected through closed-response questionnaires can provide some evidence, it may not be the best source of information for answering the questions outlined above. CT, as an emerging topic in educational and psychological research, is not well enough understood to inform the design of close-response questions that delve into the problem-solving processes directly (Leighton, 2017, p. 5; Liu et al., 2021). The limitations of closed-response questionnaires as well as numerical data warrant a data collection method that allows a wide range of responses and contributes to an understanding of problem-solving processes. Think-aloud interviews, a type of cognitive assessment tool which enables real-time data collection through verbalization, can satisfy this need.

Think-aloud interviews have grown in popularity as tools for collecting real-time data about participants' cognitive processes across a range of subject domains, such as computer science, mathematics, and clinical reasoning (Desoete, 2008; Hu & Gao, 2017; Lajoie et al., 2019). Particularly, think-aloud verbalization has been suggested as a valuable data collection method to investigate learners' cognitive processes during CT-related activities (Liu et al., 2021). However, although think alouds are often recommended in reviews of CT research (Lye & Koh, 2014; Tang et al., 2020), there is a lack of systematic understanding about their specific roles in CT studies.

Research Purpose

The purpose of this review is to examine the use of think-aloud interviews in CT studies, with a focus on the CT-related contexts in which think-aloud interviews are conducted and on the ways in which think-aloud interviews are incorporated into CT-related activities. Specifically, the following research questions are posed:

1. What are the characteristics of the CT studies that include think-aloud interviews?
2. How are think-aloud interviews incorporated into CT studies?

Research Contributions

This study investigates the application of think-aloud interviews in CT studies, highlights the potential benefits and challenges associated with this methodology, and provides valuable insights for future research. By identifying existing limitations and offering practical suggestions, this study not only advances the understanding of CT problem-solving processes, but also serves as a reference for researchers aiming to incorporate think-aloud interviews in their studies, ultimately contributing to more rigorous, inclusive, and impactful research in the field of CT and educational psychology.

Organization

Chapter 1 introduces the study and provides an overview of the context, research questions, and significance of the study. Chapter 2 presents the guiding framework for the study. Chapter 3 reviews the relevant literature on CT and think-aloud interviews. Chapter 4 shows the details of the methods adopted in this study. Chapter 5 presents the results of the study. Chapter 6 discusses the major findings of this study in light of previous literature. Chapter 7 outlines the educational implications of the study, providing suggestions for CT researchers. Chapter 8 concludes the current study by summarizing the key findings and suggesting potential directions for future research.

Chapter Summary

Chapter 1 introduces the topic of computational thinking, as well as the problem tackled in this study and its importance for education. It introduces the research purpose and questions that guide this study, followed by the potential contributions of the current study. The chapter concludes with the organization of the document, providing an overview of the structure and flow of this document.

Chapter 2 Theoretical Framework

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020; Page et al., 2021) model serves as the guiding framework for this research. PRISMA 2020 is an evidence-based, updated guideline for conducting systematic reviews and meta-analyses, designed to improve the clarity and transparency of the research studies. The PRISMA model can be used as a basis for reporting systematic reviews with objectives of evaluating empirical studies (Moher et al., 2009), and it has been adopted in previous systematic reviews related to contemporary education (Bond, 2020; Butler-Henderson & Crawford, 2020).

The PRISMA 2020 model includes a four-phase flow diagram that guides researchers throughout the systematic review process. This flow diagram visually represents the progression of information across different stages of the review, comprising identification, screening, eligibility assessment, and final study inclusion. The PRISMA 2020 model provides a structured approach to selecting, evaluating, and synthesizing relevant literature on think-aloud interviews in contemporary CT research. By adhering to this framework, this study ensures that the review process is comprehensive, transparent, and reproducible, ultimately contributing to a higher quality of evidence synthesis and more reliable conclusions.

Chapter Summary

Chapter 2 presents the PRISMA 2020 model as the guiding framework for this review, offering a structured approach to conducting a systematic review. By employing this framework, a rigorous and consistent examination of think-aloud interviews in CT research is ensured, enhancing the overall quality of evidence synthesis and the reliability of the conclusions drawn.

Chapter 3 Literature Review

Computational Thinking (CT)

CT Definition

CT is an interdisciplinary concept in relation to computer science (CS) and human cognition. Although CT has attracted considerable attention in the field of education, there is little consensus regarding its definition. During the development of constructionist learning theories, Papert (1980, 1991) proposed CT as an educational output, and further claimed that the social and affective dimensions of CT were equally important as the technical content. Despite the fact that the emerging concept of CT could be dated back to the 1980s, the modern wave of CT studies was mainly influenced by Wing's seminal proposal in 2006. Based on computer science concepts, Wing (2006) suggested CT as an ability for everyone to master in efforts to analyze and solve problems, and further addressed the importance of integrating CT into educational practice. Although Wing did not specify the dimensions and aspects of CT in the first place, her proposal provided a brand-new perspective to understand the relationship between humans and computers, resulting in a proliferation of CT studies in the following years.

Tang et al. (2020) summarized the operational definitions of CT and classified them into two main categories. While some researchers regard CT as a set of skills related to programming and computing concepts (Brennan & Resnick, 2012; Denner et al., 2014), others define CT as a competency that requires both domain-specific knowledge and general problem-solving strategies (Selby & Woollard, 2013; Yadav et al., 2014). One of the earliest CT frameworks (CSTA & ISTE, 2011) conceptualized CT as a problem-solving process involving multiple steps from problem formulation to knowledge transfer. In the context of fostering CT with *Scratch*, a tool to help young people code, Brennan and Resnick (2012) innovatively distinguished three CT dimensions: concepts, practices, and perspectives. There are also some models combining CT with classroom instruction. For instance, Barr and Stephenson (2011) proposed a CT framework involving both problem-solving skills and dispositions according to the requirements of the K-12 curriculum. Similarly, Weintrop et al. (2016) developed a taxonomy containing four CT categories with a total of 22 CT dimensions based on specific classroom activities. In the current study, CT is defined as a cognitive competency that comprises (1) knowledge about computing concepts; (2) skills and strategies derived from computing practices; or (3) dispositions and attitudes toward problem-solving.

Cognitive Processes in CT

Since Wing (2006, 2011) claimed that CT entailed the process of formulating problems and representing solutions, the time dimension of CT has been emphasized by contemporary CT researchers. For example, Aho (2012) conceptualized CT as the thought processes formulating a problem so that the solutions can be represented as following from a series of computational steps and algorithms. Similarly, the K-12 Computer Science Framework (2016) defined CT as the thought processes used to express computational steps or algorithms that can be carried out by a computer.

Previous studies have identified several cognitive factors related to CT, such as fluid intelligence, visual processing, and working memory (Ambrosio et al., 2014). According to the Cattell-Horn-Carrol (CHC) model of intelligence (Schneider & McGrew, 2018), fluid intelligence refers to a problem-solving mental activity that includes fundamental cognitive processes such as hypothesis testing, classification, and identifying relations, which are also important components of CT (Anderson, 2016; McGrew, 2009). Visual processing generally refers to the ability to generate, store, retrieve, and transform visual images and sensations (McGrew, 2009). Since CT includes processing and manipulating code blocks and other visual aspects in problem-solving, it can be connected to visual processing (Ambrosio et al., 2014). Working memory, which refers to the maintenance and manipulation of information for problem-solving purposes, is also associated with CT in terms of skill acquisition.

Given the theoretical relationship between CT and different cognitive factors, investigations are warranted of how specific cognitive processes are involved in CT. According to a recent systematic review (Tang et al., 2020), CT has been examined primarily as a learning outcome rather than a cognitive ability, and the majority of existing CT assessments focus on programming skills. Although there are a few studies considering CT as a problem-solving strategy (Horn et al., 2016; Krutz et al., 2019; Yuen & Robbins, 2014), most of them focus on the skill development and little is known about the specific cognitive processes implicated in the production of CT.

Think-Aloud Interviews

The lack of understanding in learners' cognitive processes during CT problem-solving has ushered the use of think-aloud interviews in CT. A think-aloud interview is a type of cognitive assessment tool which enables real-time data collection through verbalization. In think-aloud sessions, participants are asked to verbalize their thought processes in real time as they actively and consciously solve a task (Greene et al., 2013). The think-aloud interview, which is developed based on the earlier introspection method, has its root in cognitive psychology (Solomon, 1995, p. 29). In line with the growing interest in how people think, the think-aloud method attracted increasing attention in the late 1960s because it can provide data about the cognitive processes (Solomon, 1995, p. 31). For example, Newell and Simon (1972) combined the think-aloud protocol with problem-solving models. They recorded participants thinking aloud as they solved puzzles and then reconstructed the problem-solving processes using the verbal reports. Nisbett and Wilson (1977) criticized the use of think-aloud interviews by highlighting participants' lack of access to higher-order cognitive processes. Responding to the above critique, Ericsson and Simon (1984, 1993) modeled thinking aloud into three levels, and argued that think-aloud interviews were useful in investigating participants' cognitive processes and would not impair the problem-solving performance as long as the appropriate prompts were employed. Building on Ericsson and Simon's (1993) work, Leighton (2017) presented the theoretical basis and empirical evidence for the potency of think-aloud interviews in psychological and educational settings. Moreover, she distinguished think-aloud interviews and cognitive labs by underlining their distinct relations to human cognition (i.e., problem-solving versus comprehension). Specifically, think-aloud interviews encapsulate participants' immediate thoughts during problem-solving, shedding light on real-time cognitive processes. In contrast, cognitive labs concentrate on participants' interpretative processes, underscoring the depth of comprehension and reflection. Table 1 provides an overview of Leighton's (2017, p. 11) distinction between think-aloud interviews and cognitive labs.

Table 1

Distinction Between Think-Aloud Interviews and Cognitive Laboratory Interviews

	Think-Aloud Interviews	**Cognitive Laboratory Interviews**
Measurement Objective	Confirm/Revise model	Generate/Confirm model
Measurement Focus	Problem-solving processes	Comprehension processes
Measurement Contents	The mental activity that takes place in working memory	The organizational structure of the information that is transferred to long-term memory
Interview Probes	Concurrent interview probes	Concurrent or retrospective interview probes
Types of Verbalizations	Type 1 and 2 verbalizations	Type 3 verbalizations
Aim of Instruction Statements	Require participants to verbalize observations of straightforward environmental stimuli or provide descriptions of immediate mental activity taking place in working memory	Require the participant to search long-term memory and hypothesize on the motivations for a given solution path

Note. Adapted from *Using Think-Aloud Interviews and Cognitive Labs in Educational Research*, by J. P. Leighton, 2017, p. 11, Oxford University Press.

Think-aloud interviews have great potential in exploring a field that lacks clear and well-agreed models of cognition. One example is expertise. de Groot (1965) used the think-aloud method to explore the expertise of world-class chess players and presented a careful analysis of the conditions of specialized learning. In this pioneering study, chess masters and less experienced chess players were asked to think aloud as they decided on the move they would make. Through comparing the cognitive processes of masters and less experienced players, De Groot (1965) concluded that differences in expertise could not be attributed entirely to differences in general strategies; he pointed out that masters were able to make superior moves because they were more likely to recognize meaningful chess patterns and realize the strategic implications of these situations. Besides chess playing, think-aloud interviews have been used to study the expertise in other domains, such as teaching and clinical reasoning. In Sabers et al.'s (1991) study about pedagogical expertise, expert and novice teachers were asked to think aloud what they were seeing while watching a videotaped science lesson. Results showed that expert teachers noticed more patterns and features than novice teachers, providing insights for improving instruction (Sabers et al., 1991). Boshuizen and Schmidt (1992) explored the role of biomedical knowledge in clinical reasoning by conducting think-aloud interviews with experts, intermediates, and novices. Findings showed that experts had more in-depth biomedical knowledge than intermediates and novices, supporting a three-stage model of expertise development in medicine (Boshuizen & Schmidt, 1992).

According to Ericsson and Simon (1993), think-aloud interviews are used to measure problem-solving processes, especially dynamic processes that involve the manipulation and transformation of information in working memory. As an active cognitive processing activity, problem-solving continuously leads to a goal-directed solution or set of solutions (Leighton, 2017, p. 25). The problem-solving processes measured using think-aloud interviews have to be controlled, so that participants are conscious of the cognitive activity and generate the solution in sequence (Leighton, 2017, p. 28). Moreover, the problem-solving processes measured in think-aloud interviews are thought to rely on fluid intelligence and working memory, which are two key cognitive factors related to CT (Ambrosio et al., 2014).

A think-aloud interview can serve as a direct tool to measure the cognitive processes implicated in CT. Notably, several empirical studies have shown the potential of think-aloud interviews to externalize learners' cognitive processes during CT problem-solving (Basu et al., 2016; Jones-Harris & Chamblee, 2017; Lee et al., 2014). In addition, think-aloud interviews are frequently suggested as necessary components of CT assessments (Lye & Koh, 2014; Tang et al., 2020). Despite the benefits of think-aloud interviews, a comprehensive understanding of their relation to CT is still lacking. Therefore, this review aims to examine the use of think-aloud interviews in the CT literature.

As the core components of interviews, verbalizations can occur either concurrently or retrospectively with solving problems. The concurrent verbalization refers to the verbalizing process during the task, whereas the retrospective verbalization refers to the verbalizing process after task completion (Ericsson & Simon, 1993; Fan et al., 2019). In think-aloud interviews, the probes that encourage participants to verbalize concurrently are recommended since they are more sensitive to the memory system, especially the working memory (Leighton, 2017, p. 87).

Chapter Summary

Chapter 3 presents a comprehensive literature review on the construct of CT, cognitive processes in CT, and think-aloud interviews. The first part of this chapter delves into the definition of CT, its multifaceted nature, and various frameworks to understand CT as a cognitive competency. The second part of this chapter reviews the historical development of think-aloud methods and the application of think-aloud interviews in examining individuals' cognitive processes. Overall, this chapter provides a foundation for the systematic review on the use of think-aloud interviews in CT research.

Chapter 4 Methods

The searches were performed with the Boolean expression of "computational thinking" AND "think-aloud". Six databases were involved in the searching process, including Scopus, Web of Science, PsycINFO, Education Resources Information Center (ERIC), ACM Digital Library, and IEEE Xplore. To ensure the reliability and validity of this systematic review, only peer-reviewed publications were selected for further analysis. Moreover, since one of the objectives of this review is to examine the effectiveness of think-aloud interviews in elucidating the cognitive processes during CT problem-solving, only empirical studies were included in the current study.

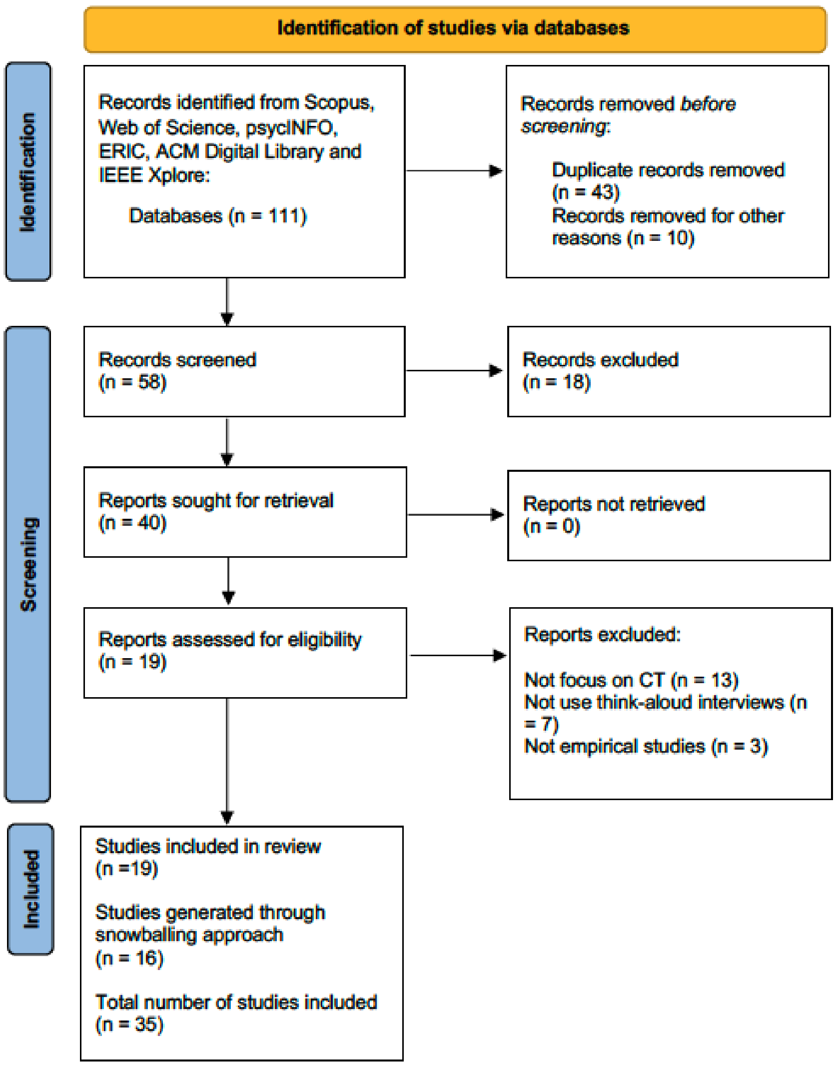

Figure 1. *PRISMA 2020 Flow Diagram*

The search and screening processes are shown in Figure 1. During the identification phase, electronic databases were systematically searched using the keywords. This was followed by the screening phase, where article titles and abstracts were assessed for relevance. Finally, in the inclusion phase, the full texts of the shortlisted articles were meticulously examined against specific criteria, ensuring only the most pertinent articles related to think alouds in CT research were incorporated into the review. With the filters of (1) language: English; and (2) document type: peer-reviewed publications based on empirical studies, the keyword searches led to a total of 111 articles, as shown in Table 2.

Table 2

The Initial Search Results

Database	N
Scopus	6
Web of Science	10
PsycINFO	2
ERIC	3
ACM Digital Library	81
IEEE Xplore	9
Total	111

The selection process consisted of multiple rounds of screening. The inclusion and exclusion criteria are presented in Table 3. The first round of screening aimed to delete duplicates, reviews, and technical articles. This process started with 111 articles and ended with 40. The second round of screening excluded 21 articles based on their keywords, abstracts, and methods. The main reasons for exclusion were a lack of (1) information about CT (n = 13); (2) details about think-aloud interviews (n = 7); and (3) an empirical focus (n = 3). At the end of a second-round screening, 19 articles remained and were further analyzed. Besides keyword searches, a snowballing approach was adopted to explore relevant articles. Through screening the bibliographies and citations of the selected articles, 16 more publications were identified and included in the current review.

During the screening of the 35 articles, the following information was extracted: (1) publication type; (2) publication year; (3) participant country; (4) grade level; (5) sample size; (6) subject domain; (7) educational setting; (8) CT definition; (9) CT-related activities; (10) CT assessment; (11) purpose of think-aloud interviews; (12) output of think-aloud interviews; and (13) reliability/validity evidence about think-aloud interviews.

All 35 articles and their extracted information were systematically organized into a comprehensive digital spreadsheet. During the first round of coding, a comprehensive approach was adopted. Each article was meticulously analyzed, with key parameters and attributes being documented. This approach sought to capture the richness of the dataset and to account for the vast diversity in study designs, methodologies, and outcomes. In the rounds that followed, the coding process underwent iterative refinement. The vast array of codes from the initial round was evaluated for clarity and consistency. Ambiguities identified were resolved, ensuring that each code was distinct and unambiguous. Redundancies or overlapping categories were consolidated to achieve a more streamlined dataset. While maintaining the depth of the first round, the aim of the refining stages was to enhance coherence and ease of analysis. The end goal was to create a coding framework that not only captured the richness of the data but also allowed for clear patterns and trends to emerge in subsequent analyses. Once coding was completed, the data underwent quantitative analyses, including frequency distributions and cross-tabulations. To ensure coding accuracy, a subset of the articles was randomly selected and reviewed. This step was essential to guarantee that consistency and reliability were maintained throughout the dataset.

Table 3

Inclusion and Exclusion Criteria

Inclusion	Exclusion
Language: English	Language: Non-English
Peer-reviewed publications	Not peer-reviewed publications

Empirical studies	Reviews, meta-analyses, and technical reports
Sufficient information related to CT	Not focus on CT
Use think alouds in the study	Not use think alouds in the study
Sufficient details of the think-aloud procedure	No details of the think-aloud procedure

Chapter Summary

Chapter 4 describes the research methodology employed in this study, encompassing the databases involved, searching keywords, inclusion and exclusion criteria, and extracted information for the systematic review. This chapter provides a comprehensive account of the methodological approach, ensuring that the study process is transparent, rigorous, and replicable.

Chapter 5 Results

What are the characteristics of the CT studies that include think-aloud interviews?

The selected 35 publications consisted of 20 conference proceedings, 12 journal articles, and three peer-reviewed book chapters. The results showed that, since the publication of the first CT study that employed a think-aloud interview in 2011, the number of publications has steadily increased during the following ten years (Figure 2). The findings revealed that 2015 was a significant time point, after which think-aloud interviews were adopted more frequently in CT research.

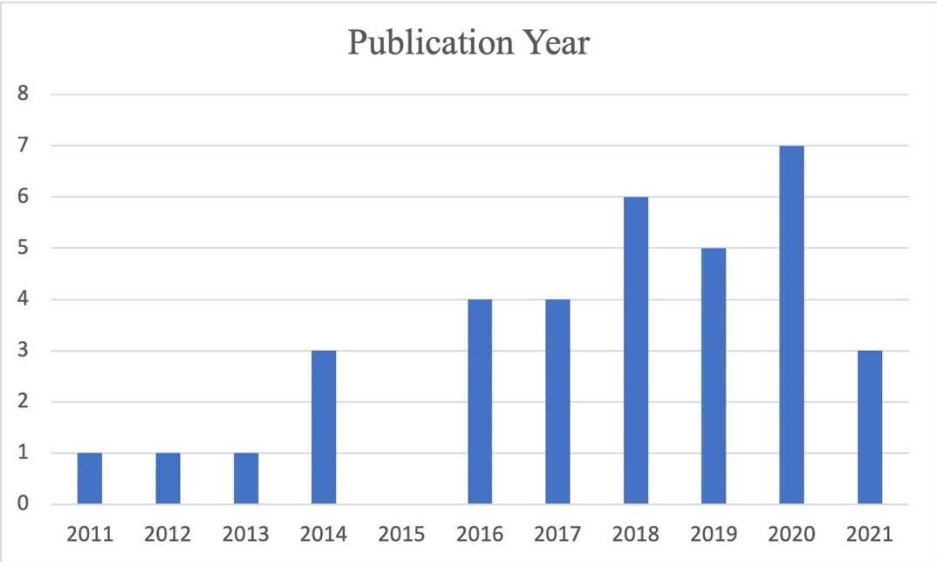

Figure 2. *Distribution of the Selected Articles by Publication Year*

Most participants were sampled from North America, specifically from the United States (n = 21), followed by Europe (n = 10) and Asia (n = 4), as shown in Figure 3. Moreover, the most representative European country was Greece (n = 6). Think-aloud interviews have been used in studies of CT across different grade levels, especially in secondary (n = 18) and elementary (n = 7) schools (Figure 4). Additionally, four studies sampled participants from both elementary and secondary schools.

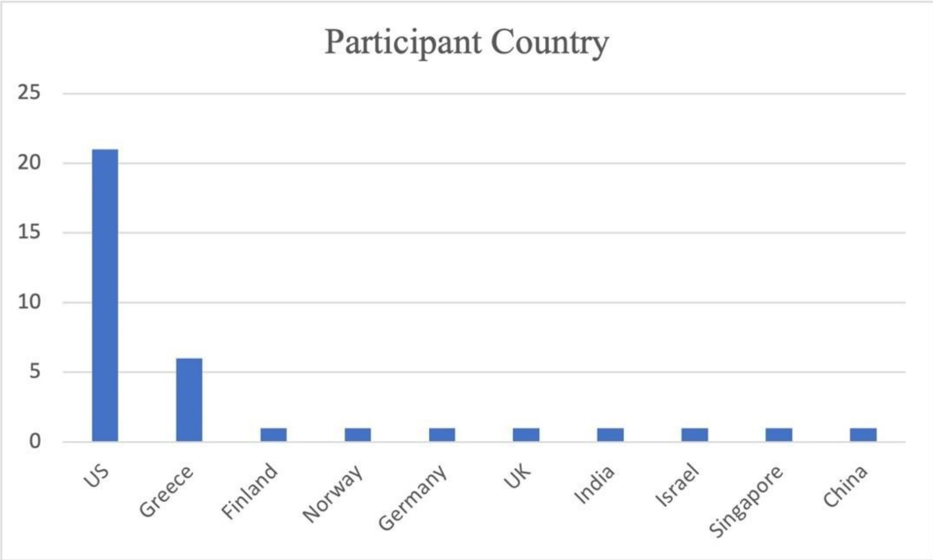

Figure 3. *Distribution of the Selected Articles by Geographical Location*

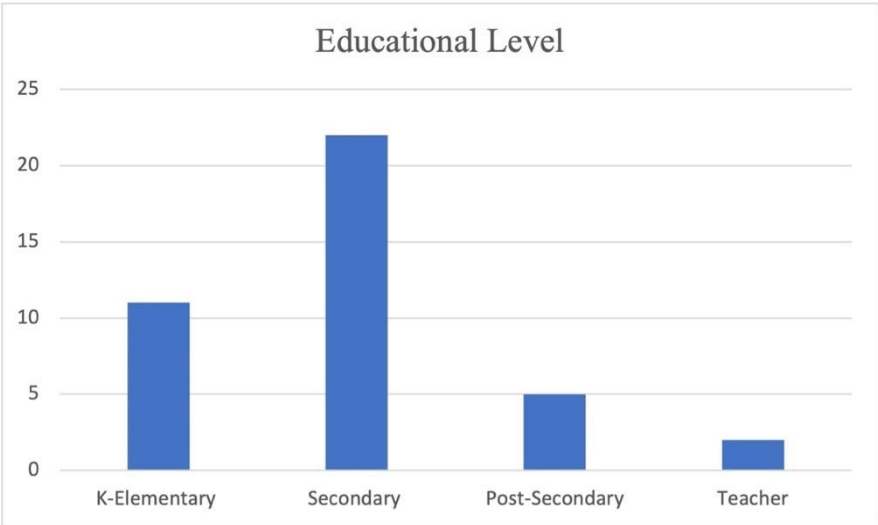

Figure 4. *Distribution of Participants' Educational Level*

More than half of the studies (n = 22) reported a sample size under 20 (Figure 5). While five studies had a sample size from 21 to 50, three studies reported a sample size larger than 50. The majority of the studies (n = 30) included some form of CS components (Figure 6). Physics (n = 5) was the second most frequent subject domain in which CT and think-aloud interviews were used, followed by robotics (n = 4) and math (n = 3). When CS was a domain of interest for conducting think-aloud interviews, it was usually combined with other subject knowledge, such as physics (n = 4) and math (n = 2).

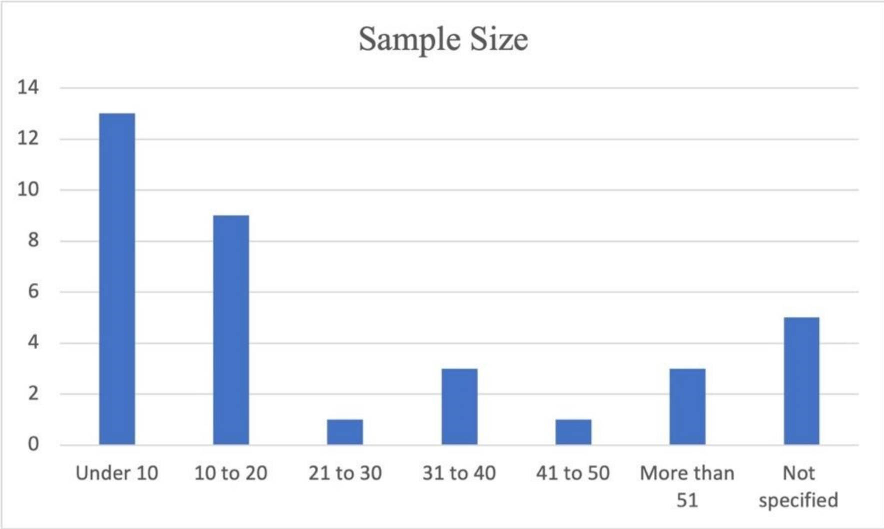

Figure 5. *Distribution of the Selected Articles by Sample Size*

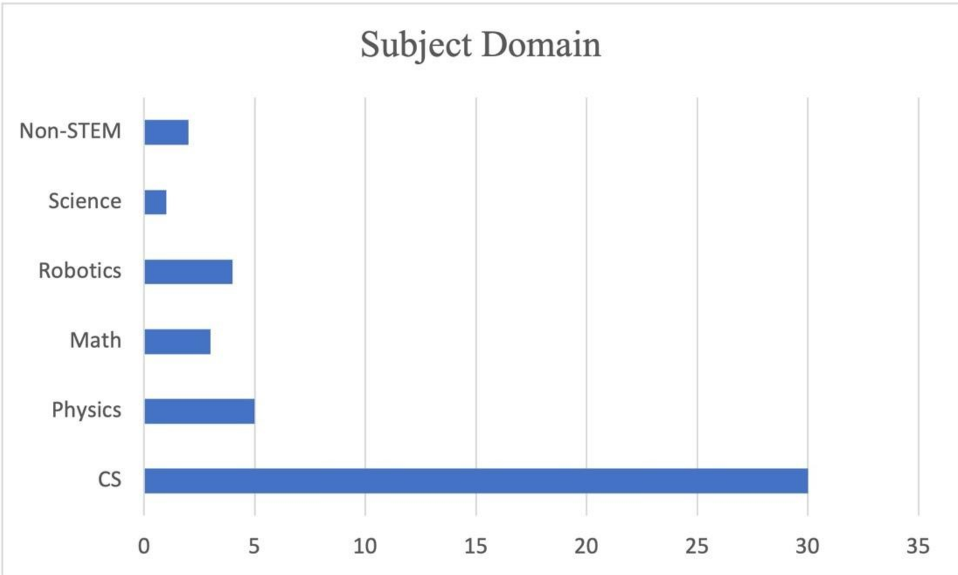

Figure 6. *Subject Domain of the Selected Articles*

As shown in Figure 7, 14 of 35 studies included programming tasks. Think-aloud CT interviews were also conducted concurrently with game play (n = 7), robotics activities (n = 4), maker activities (e.g., textile circuits, LEGO, 3D printers; n = 4), and unplugged activities (i.e., activities that do not involve a computing device; n = 4). Regarding educational settings, think-aloud interviews were more likely to be employed in formal educational settings (n = 24), which typically occurred in classrooms and during school hours.

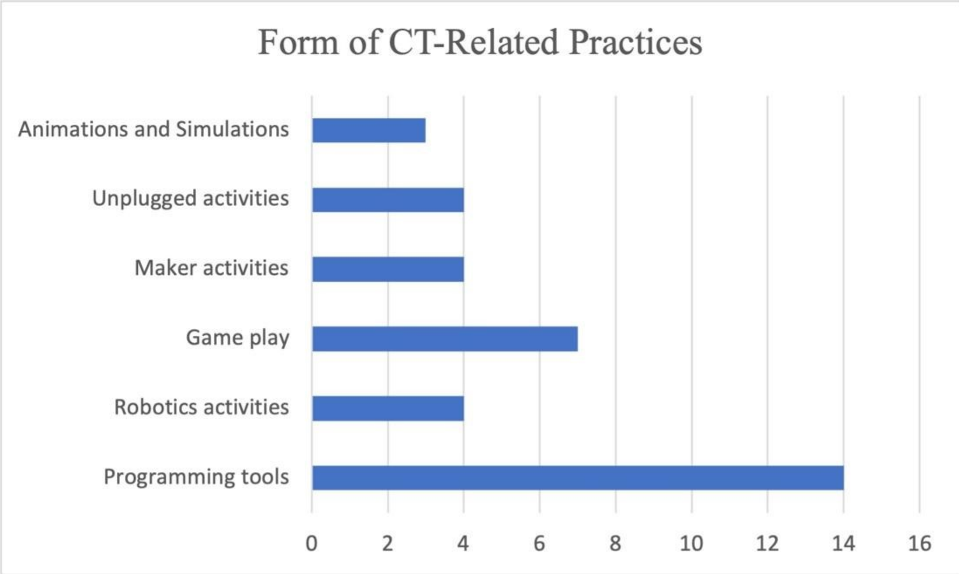

Figure 7. *Form of CT-Related Activity in the Selected Articles*

How are think-aloud interviews incorporated into CT studies?

This review involved investigating the role of think-aloud interviews in CT studies from four perspectives: (1) the timing and contexts of implementation for think-aloud interviews; (2) the purposes of conducting think-aloud interviews; (3) the presentation of the outputs of think-aloud interviews; and (4) the reliability and validity evidence.

All of the think-aloud interviews in the selected articles were conducted concurrently with the CT-related activities, such as programming tasks (n = 9), game play (n = 5), and maker activities (n = 4). In 34 of the 35 studies, the think-aloud interviews were conducted individually, enabling the investigation of individual cognitive processes.

In the studies reviewed, think-aloud interviews were seldom used as the sole assessment tool for CT. They were usually combined with other assessment tools, such as knowledge tests (n = 12) and questionnaires (n = 8), providing convergent evidence of participants' CT competency and proficiency. Generally, some of the reasons mentioned for conducting think-aloud interviews were to explore, explain, or validate. The CT studies that used think-aloud interviews for exploratory purposes (n = 19) often served as the initial phase in the investigation of participants' cognitive processes and tended to focus on CT itself, seldom connecting it with other cognitive constructs. The CT studies that used think-aloud interviews for explanatory purposes (n = 4) focused on cognitive changes rather than on specific CT concepts and practices. Such studies tended to connect CT with other cognitive constructs, such as self-regulated learning and mathematical thinking. These studies often included measurement of the sequences and transitions of cognitive strategies, and usually had clear cut-off points to determine participants' CT skill levels. Finally, the CT studies that used think-aloud interviews for validation purposes (n = 4) typically adopted more than one assessment tool to measure CT. These studies used the outputs of think-aloud interviews as supplementary resources that offered support for claims and inferences about participants' CT skill levels or developmental trajectories. Notably, eight studies conducted think-aloud interviews for multiple purposes: participants' verbal reports were not only used for validating quantitative results but also in gaining further in-depth CT insights.

The outputs of think-aloud interviews are usually presented in three forms in CT studies: quantitative analysis of code counts (n = 7); descriptive examples of think-aloud statements (n = 9); and researchers' or observers' summary of the main findings (n = 3). There are also six studies in which the outputs of think-aloud interviews were presented in a combination of two or three forms. In another ten articles, the analysis and details of participants' verbal reports were not available. In 14 of the 35 articles, a coding scheme was established before analyzing participants' verbal reports. Most of the coding schemes (n = 11) focused on CT concepts and practices, which are the components of the three-dimensional CT framework (Brennan & Resnick, 2012).

The studies that adopted more than one assessment tool for evaluating performance often used the content of think-aloud interviews as evidence for convergent validity. During the analysis of think-aloud data, eight studies presented inter-rater reliability evidence. Additionally, five studies indicated the reliability and validity evidence for the other assessment tools but not for the think-aloud interviews. Table 4 presents the basic information of the selected articles.

Table 4

Basic Information of the Selected Articles

Publication	Publication Type	Country	Grade Level	Sample Size	Subject Domain	Educational Setting	CT-Related Activities	CT Assessment	Purpose of Think-Aloud Interviews
Aiken et al., 2013	C	USA	Secondary	6	Physics	Formal	Programming in Scratch	Programming assignment, written essay, and think-aloud interviews	Exploratory
Atmatzidou & Demetriadis, 2014	J	Greece	Secondary	35	Robotics	Formal	Robotics activities	Questionnaires and think-aloud interviews	Explanatory, Validation
Atmatzidou & Demetriadis, 2016	C	Greece	Secondary	164	Robotics	Formal	Robotics activities	Questionnaires and think-aloud interviews	Exploratory
Basu et al., 2016	J	USA	Secondary	15	CS, Physics	Formal	Visual programming, simulation, and modelling activities	Knowledge tests and think-aloud interviews	Exploratory
Bonner et al., 2021	J	USA	Elementary	7	CS	Formal	Computer game play	Think-aloud interviews	Explanatory, Validation
Febrian et al., 2018	C	USA	College	5	General problem-solving	Informal	Unplugged problem-solving: 3 non-programming tasks	Think-aloud interviews	Exploratory
Fields et al., 2020	C	USA	Secondary	8	CS, Physics	Formal	Maker activities	Think-aloud sessions	Exploratory, Validation
Friend & Cutler, 2013	C	USA	Secondary	6	CS	Informal	Unplugged problem-solving: egg-drop contraptions	Think-aloud interviews and cognitive laboratory interviews	Exploratory
Gardeli & Vosinakis, 2017	J	Greece	K-12	60	CS	Formal	Visual programming in Scratch and unplugged activities	Observation, error log and analysis, think-aloud interviews, evaluation and timing of solutions, questionnaires, and open discussion	Exploratory
Gardeli & Vosinakis, 2020	B	Greece	Elementary	18	CS	Formal	AR animation, game-based programming activities	Knowledge tests; observation, think- aloud interviews, questionnaire	Exploratory, Validation

								s, and open discussion	
Grover & Basu, 2017	C	USA	Secondary	3	CS	Formal	Programming in Scratch	Knowledge tests	Validation
Horn et al., 2016	C	USA	Secondary	9	CS, Science	Informal	Computer game play	Knowledge tests, game log file, think-aloud interviews	Exploratory, Explanatory, Validation
India et al., 2020	C	India	Elementary	12	CS, Music	Formal	Toy play	Open-ended diary entries, observations, think-aloud interviews	Exploratory, Validation
Jayathirtha et al., 2020	C	USA	Secondary	14	CS, Physics	Formal	Debug a researcher-designed project	Think-aloud debugging sessions, and debugging performance	Exploratory
Jones-Harris & Chamblee, 2017	B	USA	Secondary	6	CS, Math	Formal	Unplugged activities and programming in Java	Problem-solving solutions and strategies, think-aloud and retrospective interviews, and analytic scoring rubric scale scores	Explanatory
Kluge et al., 2019	C	Norway	Secondary	37	CS	Formal	Programming tasks	Think-aloud interviews	Exploratory
Kruskopf et al., 2019	C	Finland	Elementary	14	CS	Informal	Visual programming (Computational thinking mine-task), and game play	Self-report questionnaires and think-aloud interviews	Explanatory
Krutz et al., 2019	J	USA	In-service teachers	5	CS	Informal	Two block-based programming tasks	Think-aloud sessions, task completed, and task duration	Exploratory
Lee et al., 2014	J	USA	K-12	18	CS	Informal	Social game play	Interviews and a think-aloud protocol	Exploratory
Liu & Rojas, 2019	C	China	Elementary	15	Robotics, CS	Formal	Robotics-based programming and maker activities	Questionnaires, think-aloud protocols, observations, project portfolios, surveys, pre and post-tests	Exploratory
L'Heureux et al., 2012	C	USA	College	N/A	CS	Formal	IT problem-solving projects	Survey, interviews, knowledge tests	Exploratory
Looi et al., 2018	J	Singapore	Secondary	35	CS	Formal	Unplugged activities	Audio and video recordings of the classroom activities, observation notes,	Exploratory

								worksheets, think-aloud sessions	
Luo et al., 2020	C	USA	Elementary	13	CS, Math	Formal	Programming tasks	Think-aloud sessions and test items	Exploratory
Mardi, 2020	J	USA	Graduate students (pre-service and in-service teachers)	N/A	Math	Formal	A math solution design project	Think alouds and digital powerups	Exploratory
McGee et al., 2018	C	USA	Secondary	N/A	CS	Formal	Activities that are designed to engage students in CS inquiry around meaningful problems	Knowledge tests and think-aloud interviews	Validation
Pasternak, 2016	C	German	Secondary	N/A	CS	Formal	Learn CS items in the context of everyday life	Questionnaires and think-aloud interviews	Explanatory, Validation
Richard et al., 2018	C	USA	K-12	21	CS, Physics	Informal	Maker activities	Video-taped observations, think-aloud protocols, pre- and post-surveys	Exploratory
Rowe et al., 2021	J	USA	K-12	N/A	CS	Informal	Computer game play	Think-aloud CT gameplay sessions and log data	Validation
Seo & Richard, 2018	C	USA	Secondary, Young adults	5	Robotics, CS	Informal	Robotics-based maker activities	Microanalytic video analysis, open-response pre and post questionnaires, focus group interviews, think-aloud interviews	Exploratory, Validation
Snow et al., 2017	C	USA	Secondary	8	CS	Formal	CT problem-solving	Knowledge tests and cognitive think-aloud interviews	Validation
Starbird & Palen, 2011	C	USA	Pre-college or early college	6	CS	Informal	Programming tasks, 3D animation	Computer interaction activities and think alouds	Exploratory
Statter & Armoni, 2020	J	Israel	Secondary	45	CS	Formal	Programming in Scratch	Pre- and post-tests, final projects, think-aloud interviews, and class observations	Explanatory

Turchi et al., 2019	J	UK	Secondary	18	CS	Informal	Computer game play	Think-aloud feedback, participants interactions recording and observation, and a post-test survey	Exploratory
Vourletsis et al., 2021	B	Greece	Elementary	86	CS	Formal	Debugging activities	Think-aloud protocols. Prompted written descriptions, semi-structured interviews, screen recordings and a rubric	Exploratory
Yuen & Robbins, 2014	J	Greece	Undergraduate	5	CS (for Biology major)	Formal	Programming projects in MATLAB	Pre- and post-tests, and think-aloud interviews	Explanatory

**Publication Type: C – Conference Proceeding, J – Journal Article, B – Book Chapter. N/A: Not Available.*

Table 5 outlines the research design of the selected articles, highlighting their research questions and the theoretical frameworks underpinning them.

Table 5

Research Design, Research Questions, and Theoretical Framework of the Selected Articles

Publication	Research Design	Research Questions	Theoretical Framework
Aiken et al., 2013	Mixed-methods	N/A	N/A
Atmatzidou & Demetriadis, 2014	Mixed-methods	Are students of different age and gender developing CT skills in the same way in the context of educational robotics activity?	Five core dimensions of the broader CT conceptual framework: abstraction, generalisation, algorithm, modularity and decomposition
Atmatzidou & Demetriadis, 2016	Mixed-methods	(1) How can the CT and problem-solving skills be supported efficiently in educational robotics activities? (2) Which are the appropriate strategies for assessing the development of CT?	Basic skills of CT: abstraction, generalization, algorithm, modularity, decomposition and problem solving
Basu et al., 2016	Quasi-experimental	(1) What are the different types of challenges that students face while working on CTSiM, and what kinds of supports can help them overcome these challenges? (2) How do these challenges evolve across a sequence of curricular units taking into account that students are scaffolded one-on-one by researchers when they have difficulties?	Four broad challenge categories: (1) programming challenges, (2) modeling challenges, (3) domain challenges, and (4) agent-based reasoning challenges
Bonner et al., 2021	Descriptive	(1) Do student think alouds provide evidence that computing problems in the FA task elicit cognitive processes consistent with conceptual definitions of CT? (2) Do student think alouds provide evidence that SRL prompts embedded in the FA task elicit cognitive processes consistent with SRL theory? (3) How do students at different levels of mastery vary in use of CT and SRL? (4) Do think alouds	CT practices and self-regulated learning

		provide evidence that elucidates definitions of the CT construct or SRL processes?	
Febrian et al., 2018	Qualitative	(1) In what ways do students use CT skills when solving non-programming problems if any? (2) If students use CT when solving non-programming problems, in what ways do their approaches differ from computer science students?	Grover & Pea's (2018) CT framework
Fields et al., 2020	Qualitative	(1) What were affordances and limitations of the reconstruction kit for students' debugging and designing e-textiles? (2) How were students' collaborations supported in debugging and designing e-textiles?	Rapid designing and debugging in circuitry; collaborative debugging and designing
Friend & Cutler, 2013	Qualitative	N/A	Algorithmic efficiency
Gardeli & Vosinakis, 2017	Qualitative	N/A	CT concepts (considering only logical mistakes)
Gardeli & Vosinakis, 2020	Experimental	(1) What is the perceived value and effectiveness of using a tangible interface in an AR-based programming activity compared to an unplugged activity, in terms of programming performance and motivation, then (b) how the role of the educator is differentiated in class settings concerning those two approaches, in terms of involvement and engagement and finally, (c) does the AR-based approach support the educational goals, while addressing the limitations of the current formal education settings, such as educators' skills and training, and institution resources? What are the factors and the characteristics of the system that contribute to that?	Perception concerning enjoyment and ease of use: surprise, excitement, and concerns
Grover & Basu, 2017	Qualitative	(1) How can learning outcomes for computing constructs such as variables, expressions (arithmetic and logical), and loops, be organized into a structured assessment framework and measured with technical quality? (2) What do assessments aimed to measure student understanding of computing constructs such as variables, expressions and loops, tell us about student understanding and misconceptions related to these concepts in the context of block-based programming in middle school CS?	Foundational programming constructs such as variables, loops, and expressions
Horn et al., 2016	Mixed-methods	N/A	N/A
India et al., 2020	Qualitative	Is it possible for children who are blind, studying in schools for the blind in low-resource settings, to pick up concepts in computational thinking using Project Torino with reliance on play instead of structured teaching?	Three-dimension CT framework: computational concepts (sequence, thread, loop and if-then-else), computational practice (tracing and debugging) and computational perspectives (expressing and connecting)
Jayathirtha et al., 2020	Qualitative	(1) What debugging strategies do novice high school students adopt to identify and address the problems in the e-textiles projects? (2) How do they navigate the multi-representational problem spaces as they debugged these projects?	CT: debugging strategies and approaches
Jones-Harris & Chamblee, 2017	Mixed-methods	(1) What problem-solving strategies do African-American students who have taken or are jointly enrolled in precalculus and AP computer science course demonstrate	Analytical thinking processes, mathematical thinking, CT

		when solving precalculus problems? (2) What problem-solving strategies do African-American students who have taken or are jointly enrolled in precalculus and AP computer science courses demonstrate when solving computer science problems? (3) Are there relationships between the problem-solving strategies African-American students who have taken or are jointly enrolled in precalculus and AP computer science use to solve precalculus and computer science problems?	
Kluge et al., 2019	Qualitative	(1) On what level are the students able to talk about their code in this learning environment? (e.g., discuss pros and cons, alternative solutions and more) (2) In what way may the screencasts give additional information that is potentially usable for a teacher?	PRIMM framework
Kruskopf et al., 2019	Mixed-methods	(1) Did any conceptual change take place during a wireless, play oriented summer school concerning the children's (participants) ideas of technology and their computational thinking? (2) Did any parallel conceptual change take place in two other contexts, sports camp and scout camp?	Conceptual change theories
Krutz et al., 2019	Qualitative	(1) What usability issues exist in the current prototype that should be addressed before conducting classroom studies? (2) What pedagogical insights do teachers have about introducing stepwise refinement to their students using this tool?	N/A
Lee et al., 2014	Experimental	(1) Did children perceive CTArcade (COMP) as being a more enjoyable activity compared to paper (PAP)? (2) What explanations were provided for the participants' preferences?	CT skills: algorithmic thinking; pattern generalization and abstraction; problem decomposition; pattern recognition
Liu & Rojas, 2019	Experimental	N/A	CT skills which include five key concepts: abstraction, analysis, generalization, algorithmic thinking and logical thinking
L'Heureux et al., 2012	Quasi-experimental	N/A	N/A
Looi et al., 2018	Mixed-methods	How can an unplugged CT activity interplay with the production of an artifact to educed the development of CT?	Decomposition, abstraction of data and functionality, generalization, algorithmic design, and evaluation and improvement
Luo et al., 2020	Qualitative	(1) How do 4th-grade students express and articulate CT understanding, in the areas of sequence, repetition, decomposition, and conditionals? (2) How does students' CT understanding correspond to the hypothetical cognitive progression of the learning trajectories?	Four CT concepts: sequence, repetition, conditionals, and decomposition
Mardi, 2020	Qualitative	How do graduate students use empathy and CT terms as they reflect on their problem-solving tasks and projects?	CT can be visualized as a series of core concepts (logic, evaluation, algorithms, patterns, decomposition, and abstraction) and approaches (tinkering, creating, debugging, persevering, and collaborating)
McGee et al., 2018	Quasi-experimental	N/A	CT practices
Pasternak, 2016	Mixed-methods	N/A	N/A

Richard et al., 2018	Qualitative	(1) How do age-blended elementary and middle school learners demonstrate individual and collaborative learning with multimodal tools and affordances? (2) How do they engage STEM and non-STEM interests through the curriculum?	N/A
Rowe et al., 2021	Mixed-methods	(1) What indicators of implicit CT can be reliably predicted by automated detectors in Zoombinis? (2) How do in-game measures of implicit CT in Zoombinis relate to external measures of CT? (I.e., are these valid assessments?)	CT practices: problem decomposition, explicit problem decomposition, implicit problem decomposition, pattern recognition, abstraction
Seo & Richard, 2018	Qualitative	(1) How do LVIs engage with platforms that are tangibly accessible? (2) How does the accessibility of the tools affect learners' self-efficacy and collaborative interactions? (3) What design elements do LVIs express would be beneficial for equitable co-creation and collaborative learning?	Computational concepts and practices
Snow et al., 2017	Mixed-methods	N/A	N/A
Starbird & Palen, 2011	Qualitative	How novice programmers, within an environment that has no "teacher," use borrowed code in different ways to learn computer programming skills?	N/A
Statter & Armoni, 2020	Mixed-methods	How does using the framework for teaching CS abstraction in a 7th-grade introductory CS course affect the students' CS abstraction skills?	A framework for teaching abstraction in the context of algorithmic problem-solving, intended for novice students
Turchi et al., 2019	Qualitative	(1) Is gameplay an effective way to foster learning of CT skills? (1a) Can we provide a playful way of learning CT skills? (1b) Can collaborative learning help improving CT skills?	CT skills: problem-solving, algorithmic thinking, abstraction, decomposition, thinking recursively, using heuristics, data representation, evaluating solutions (debugging)
Vourletsis et al., 2021	Quasi-experimental	(1) What is the effect of a CT instructional intervention on the testing and debugging proficiency level of 6th-grade pupils? (2) What is the effect of a CT instructional intervention on the ability of 6th-grade pupils to employ systematic debugging strategies?	Three-dimension framework: concepts, practices, and perspectives
Yuen & Robbins, 2014	Qualitative	(1) How CT skills were developed? (2) How quantitative skills were developed?	(1) organization: coding style, data organization; (2) construction: following procedures, visualizing data; (3) analysis: interpretation, conclusions

Table 6 delves into the specifics of the think-aloud interviews used in the selected articles, showing the content covered, the analyses undertaken, and the coding schemes employed.

Table 6

Contents, Analyses, and Coding Scheme Related to Think-Aloud Interviews

Publication	Contents of Think-Aloud Interviews	Think-Aloud Analyses	Coding Scheme
Aiken et al., 2013	Cognitive processes and strategies used while completing the scaffolded code and answering questions about physics concepts	Descriptive examples of think-aloud statements	N/A

Atmatzidou & Demetriadis, 2014	The process followed to solve a certain robot programming task	Quantitative analysis of code counts	Yes (derived from the graded criterion of the instrument, 4-point Likert scale)
Atmatzidou & Demetriadis, 2016	The process followed while solving a robotics problem	Quantitative analysis of code counts	Yes (derived from the graded criterion of the instrument, 4-point Likert scale)
Basu et al., 2016	Cognitive processes while working with CTSiM	Quantitative analysis of code counts	Yes (the challenge and frequency counts for activities)
Bonner et al., 2021	Cognitive and metacognitive processes in formative assessment tasks during game play	Researchers' or observers' summary of findings (profile); descriptive examples of think-aloud statements	Yes (theory-driven)
Febrian et al., 2018	Cognitive processes while solving non-programming problems	Descriptive examples of think-aloud statements (whether a specific CT skill is present); researchers' or observers' summary of findings	Yes
Fields et al., 2020	Working processes while working on the designing and debugging e-textiles projects	Researchers' or observers' summary of findings	N/A
Friend & Cutler, 2013	Thought processes while solving the problem of egg-drop contraptions ("How would you solve this problem?"; "Which of these three solutions do you think is the best? Why?")	Descriptive examples of think-aloud statements; researchers' or observers' summary of findings	N/A (open-coding protocol)
Gardeli & Vosinakis, 2017	Thought processes when testing the algorithms in an unplugged and a visual programming environment	Descriptive examples of think-aloud statements	N/A
Gardeli & Vosinakis, 2020	Cognitive processes while creating three programming solutions in a game-based learning activity	Descriptive examples of think-aloud statements	N/A
Grover & Basu, 2017	Thought processes while working on assessment items in Scratch	Descriptive examples of think-aloud statements	N/A
Horn et al., 2016	Cognitive processes while game playing	Quantitative analysis of code counts	Yes
India et al., 2020	Thoughts processes during building the program	Researchers' or observers' summary of findings	N/A
Jayathirtha et al., 2020	Thought processes and interactions while debugging the Debuglts	Quantitative analysis of code counts	Yes (distributed cognition)
Jones-Harris & Chamblee, 2017	Metacognitive processes and analytic thinking processes	Quantitative analysis of code counts	Yes (analytic scoring rubric scale)
Kluge et al., 2019	Working processes during coding	Descriptive examples of think-aloud statements	N/A
Kruskopf et al., 2019	Thought processes during game play	Descriptive examples of think-aloud statements	Yes
Krutz et al., 2019	Explaining the actions when completing the programming task	N/A	N/A
Lee et al., 2014	Thought processes during each move during game play	Quantitative analysis of code counts; descriptive examples of think-aloud statements	Yes
Liu & Rojas, 2019	Working processes to solve the robot programming task	N/A	N/A
L'Heureux et al., 2012	N/A	N/A	N/A
Looi et al., 2018	Strategies of using a balance scale to sort eight unknown weights in ascending order from the lightest to the heaviest	N/A	N/A
Luo et al., 2020	Thought processes when addressing each assessment item	Quantitative analysis of code counts	Yes

Mardi, 2020	Cognitive processes while reflecting on a math solution design project	Descriptive examples of think-aloud statements (content analysis)	Yes (codes from the literature)
McGee et al., 2018	N/A	N/A	N/A
Pasternak, 2016	Thought processes during solving the programming problems	N/A	N/A
Richard et al., 2018	Metacognitive processes during problem-solving in a maker workshop	Descriptive examples of think-aloud statements (interaction analysis)	N/A
Rowe et al., 2021	Thought processes as solving the puzzles during game play	N/A	N/A
Seo & Richard, 2018	Cognitive processes and interactions during the maker activities	N/A	N/A
Snow et al., 2017	Read the scenario out loud and talked through the solution	Researchers' or observers' summary of findings	N/A
Starbird & Palen, 2011	Thought processes ("What are you thinking about?")	Descriptive examples of think-aloud statements; researchers' or observers' summary of findings	Yes (a qualitative analytical coding scheme)
Statter & Armoni, 2020	Thought processes during programming tasks in Scratch	Descriptive examples of think-aloud statements	N/A
Turchi et al., 2019	Thought processes during game play	N/A	N/A
Vourletsis et al., 2021	Whatever they saw, processed, performed, and felt as performing the debugging tasks	N/A	N/A
Yuen & Robbins, 2014	(1) "What are you doing?" and "Why are you doing this?" (2) When participants received feedback from their programs, they were asked to explain the feedback they received. (3) The interviewer probed participants' understanding by asking them to explain the concepts involved in their task—both data analysis concepts and computing concepts. (4) The participants were asked to discuss or experiment with different scenarios, coding implementations, and alternate solutions	Descriptive examples of think-aloud statements; quantitative analysis of code counts	Yes (open-coding, axial coding, and selective coding)

Table 7 underscores the contributions of the think-aloud interviews, alongside evidence related to their reliability and validity in the context of the studies presented.

Table 7

Contributions and Reliability/Validity Evidence of Think-Aloud Interviews

Publication	Contribution of Think-Aloud Interviews	Reliability/Validity Evidence
Aiken et al., 2013	Investigated how students make connections between physics concepts and CT	N/A
Atmatzidou & Demetriadis, 2014	Identified students' CT skills levels in the context of educational robotics (ER) learning activity	N/A (the reliability/validity evidence was about the whole study instead of think-aloud interviews)
Atmatzidou & Demetriadis, 2016	Saw the progression of students' CT skills and attitudes in Educational Robotics activities	N/A
Basu et al., 2016	Identifies the challenges pertaining to the processes students employed when constructing simulation models in CTSiM to learn about topics and concepts in kinematics and ecology	Reliability with the research codes; inter-rater reliability

Bonner et al., 2021	Illustrated the length and sequence of different types of cognitive processes involved in performance of a formative assessment of CT designed to support self-regulated learning	Inter-rater reliability; content validation approach
Febrian et al., 2018	Identified various CT skills; provided insights on whether people used CT skills when solving non-programming problems	Inter-rater reliability
Fields et al., 2020	Investigated the potential affordances of a revised reconstruction kit for e-textiles that would allow more rapid collaborative prototyping and debugging of e-textile circuitry and designs	N/A
Friend & Cutler, 2013	Categorized solutions that students chose and the CT skills they used in solving the problem of egg-drop contraptions	N/A
Gardeli & Vosinakis, 2017	Provided insights into students' errors	N/A
Gardeli & Vosinakis, 2020	Provided further evidence into the effect of an AR-based learning interface on student motivation, effectiveness, and teaching practice	N/A
Grover & Basu, 2017	Examined students' responses and misconceptions about the assessment items that measured student understanding in middle-school introductory CS classrooms	N/A (the evidence of inter-rater reliability was about the whole study instead of think-aloud interviews)
Horn et al., 2016	Provided insights into players' in-the-moment strategies, perceptions, and goals	Convergent evidence
India et al., 2020	Investigated how students were approaching each step and what pods they were going to use	N/A
Jayathirtha et al., 2020	Expanded the understanding of debugging in the physical computing systems	Inter-rater reliability; reach a consensus by discussion
Jones-Harris & Chamblee, 2017	Determined the levels of students' problem-solving ability and accessed students' analytic thinking strategies and solutions between similar mathematics and computer science tasks	Inter-rater reliability
Kluge et al., 2019	Investigated how the students explained and expanded their program code in the programming environment and determined the details of how the students used screencasts to convey content to their teacher	N/A
Kruskopf et al., 2019	Investigated the misunderstandings and incoherencies in children's CT thought processes	N/A
Krutz et al., 2019	Incorporated stepwise refinement into block programming environments as an approach for novices to work out the CT processes needed to write more complex programs	N/A
Lee et al., 2014	Saw how young children articulated CT skills under the various game play conditions	Inter-rater reliability
Liu & Rojas, 2019	Showed students' mastery of programming knowledge and CT skills	N/A
L'Heureux et al., 2012	Investigated the implementation of information technology problem-solving constructs and scenarios designed to cultivate computational thinking in information technology education at the college level via a course entitled "IT Problem Solving."	N/A
Looi et al., 2018	Explored how an unplugged sorting activity educed the development of CT skills for grade 9 students	N/A
Luo et al., 2020	Shed light into students' articulated understanding of the four CT concepts and the correspondence between that understanding and hypothesized learning trajectories	Inter-rater reliability; followed the principled design process; conducted an internal review process; statistical models fitted on student response data
Mardi, 2020	Responded to the questions: (1) Which empathy and CT terms do students use to reflect throughout the course tasks? (2) How do they relate those terms to their problem-solving process or task?	N/A
McGee et al., 2018	Provided validity evidence for patterns in the development of CT practices in the context of Exploring Computer Science (ECS) program	N/A

Pasternak, 2016	Knew if the students had perceptions linked to CS concepts	N/A
Richard et al., 2018	Elicited learners' reflections scaffolded their problem-solving	N/A
Rowe et al., 2021	Mirrored the range of conditions of classroom implementation in which the detectors will likely be used	N/A (the reliability/validity evidence was about the whole study instead of think-aloud interviews)
Seo & Richard, 2018	Advanced the understanding of how accessibility affected both individual and collaborative group interaction and cognition with the artifacts	N/A
Snow et al., 2017	Provided validity evidence for areas in which the students struggled either with the interpretation of the question or with determining how to answer it. This information was used to guide questions revisions to reduce confusion and to ensure that the tasks had no constructs that would cause variance among students that was irrelevant to the targeted knowledge and skills.	N/A (the reliability/validity evidence was about the whole study instead of think-aloud interviews)
Starbird & Palen, 2011	Investigated how non-programmers' use of a new platform for end-user programming	N/A
Statter & Armoni, 2020	Examined a few aspects including (1) conscious transition between levels of abstraction, (2) using correct algorithms and distinguishing between the algorithm level and the programming level, (3) using black boxes naturally, (4) repairing faults in the program, and (5) the perception of the learning outcomes of the course	The questions in think-aloud interviews were content-validated by an expert
Turchi et al., 2019	Provided insights on how participants responded to playing with TAPASPlay and managed to build different strategies in order to win the game	N/A
Vourletsis et al., 2021	Measured students' debugging proficiency level and strategy use on four learning units of the instructional intervention	N/A
Yuen & Robbins, 2014	Identified different processes and constructs—and how they worked together—during data-driven computing activities	Inter-rater reliability

Chapter Summary

Chapter 5 presents the results from the systematic review, examining the characteristics of CT studies that incorporate think-aloud interviews and the manner in which these interviews are integrated into the research. The findings encompass various aspects, such as the distribution of publications over time, demographic information of participants, subject domains in which CT is studied, CT-related activities, and the educational contexts where the research takes place. Additionally, this chapter investigates the distinct objectives behind employing think-aloud interviews in CT studies, including exploratory, explanatory, and validation goals. It elucidates the approaches employed to present and evaluate the results of think-aloud interviews in the analyzed studies.

Chapter 6 Discussion

Integrating CT in contemporary education can not only facilitate students' cognitive development (Lockwood & Mooney, 2018), but also improve their learning motivation toward CS (Allan et al., 2010) and other STEM subjects (Sengupta et al., 2013). Although extensive research has been conducted on CT, little attention has been paid to the associated cognitive processes that lead to successful CT. As suggested by Kallia et al. (2021), the cognitive processes used in problem-solving can be regarded as the core processes of CT, and they need to be assessed technically. A think-aloud interview is a real-time data collection method that has been proposed to help elucidate the processes that permit students to engage in CT. However, despite the call by several research reviews to empirically uncover the cognitive processes involved in CT (Lye & Koh, 2014; Tang et al., 2020), little is known about their specific role in CT.

The present review adopted a systematic framework to investigate the use of think-aloud interviews in CT studies. Generally speaking, think-aloud interviews have several advantages over traditional assessments. On the one hand, the method of verbal data collection may involve more diverse participants, such as visually impaired children in developing countries (India et al., 2020). On the other hand, qualitative data have the potential to help researchers generate deeper insights about CT, especially in terms of the cognitive processes involved.

What are the characteristics of the CT studies that include think-aloud interviews?

This review summarized the characteristics of CT studies in which think-aloud interviews were used. Although the resurgence of CT research started in 2006, the think-aloud interview was not considered as an assessment tool for CT abilities until 2011.

Additionally, the think-aloud interview was more likely to be employed in CT studies after 2015, and such a trend may have been influenced by the suggestion made in a relevant systematic review by Lye and Koh (2014). Regarding the geographical distribution of the studies, most CT-related think-aloud studies were conducted in the United States. This finding is not surprising as most CT research is conducted in the United States (Cutumisu et al., 2019; Hsu et al., 2018).

Another salient feature of the selected CT studies is the relatively small sample sizes used. Compared with knowledge tests and questionnaire surveys, think-aloud interviews are more time-consuming and labor-intensive to conduct, resulting in the restriction of sample size to some extent (Cotton & Gresty, 2006; Johnstone et al., 2006; Merchie & Van Keer, 2014). However, one of the missions of qualitative studies is to gather extensive data from a small sample in order to generate richer and in-depth conclusions (Fonteyn et al., 1993). Generalization of results to a wider population is often not the goal of qualitative studies (Creswell, 2013). As stated by Leighton (2017, p. 36), the sample size chosen is subject to specific considerations and depends on the objective of the investigation. In most CT studies, think-aloud interviews are used for exploratory purposes. In this sense, small sample size may represent a good balance between the cost and the thoroughness of the investigation. Nevertheless, the small sample size does not indicate small amounts of data (Johnstone et al., 2006). By means of protocol analysis, the data generated from think-aloud interviews can be extensive.

Regarding the educational context in which the studies were conducted, most of the selected studies were conducted in formal settings and related to CS and programming, consistent with the conclusions of previous reviews (Lockwood & Mooney, 2018; Tang et al., 2020). Particularly, the higher proportion of studies conducted in formal educational settings may reflect the increasing involvement of CT in the standard curriculum. The findings also revealed that most of the studies were conducted in K-12 settings, with few studies if any including adult learners. Consistent with findings of recent reviews (Cutumisu et al., 2019;

Tang et al., 2020), most CT interventions as well as assessment tools are mainly designed for K-12 learners, possibly because CT interventions may be integrated within a standard curriculum and during school time. However, there are possible drawbacks to the singular focus on K-12 students: in think-aloud sessions, young learners are likely to be distracted and may not always be aware of their cognitive processes (Mueller et al., 2017). Therefore, they may not be able to demonstrate their CT problem-solving processes through verbalization comprehensively. Thus, think-aloud interviews may reveal more when applied to adults compared with young learners. To develop a complete picture of the cognitive processes used in CT, studies of adult participants should be conducted. Continuing with the settings in which the studies were conducted, secondary CS classrooms were the most typical context for conducting think-aloud interviews.

Although most of the CT-related activities presented to participants were computer-based, some activities were unplugged. Recently, the emphasis of CS education has shifted from learning how to program in a specific language towards developing CT more generally in terms of problem-solving skills (Guzdial, 2008). The popularity of unplugged activities that take place away from digital devices (Brackmann et al., 2017; Rodriguez et al., 2017) makes it easier to disassociate CT from computer-based environments (Bell et al., 2009). These activities typically use games, cards, strings, and physical movements to demonstrate computing concepts (e.g., algorithms and decomposition; Brackmann et al., 2017). Unplugged activities offer learners the opportunity to apply CT in the solving of real-world problems (del Olmo-Muñoz et al., 2020; Looi et al., 2018) and can address the problem of unreliable or insufficient computers, involving more diverse participants, especially underrepresented populations in CT education (Huang & Looi, 2020). Similar to unplugged activities, think-aloud interviews also make it possible to assess CT in underrepresented populations. Think-aloud interviews can be combined with unplugged activities to study the cognitive processes used in CT, for example, by exceptional learners.

Finally, one area that has been relatively neglected is the potential influence of different think-aloud interview probes (e.g., concurrent verbalizations) on CT. There are some studies showing that the use of think-aloud interviews can help students improve their reading comprehension and achievement (Fisher et al., 2011; Ness, 2016). In addition to measuring problem-solving processes in CT, think-aloud procedures may be useful in improving the expression and confidence of participants (e.g., Liu & Rojas, 2019).

How are think-aloud interviews incorporated into CT studies?

This review explored the ways in which think-aloud interviews were incorporated into CT studies. As methodological tools, interviews can be implemented either concurrently or retrospectively. The concurrent interviews involve the verbalization processes during the task. During a concurrent verbalization session, participants attempt to verbalize the cognitive processes they engage in and the information they attend to while solving the problem (Leighton & Gierl, 2007). In contrast, retrospective verbalizations are always conducted in the form of cognitive interviews after task completion, enabling participants to describe and reflect on their cognitive processes (Ericsson & Simon, 1993). Ericsson and Simon (1993) further identified three sorts of verbal statements and suggested that only two of them (i.e., type 1 and 2 verbalization) could be utilized as evidence of problem-solving processes in working memory. Type 1 and 2 verbalizations require participants to verbalize their direct observations of the external stimuli or describe their immediate mental activity occurring in the working memory (Leighton, 2017, p. 41). In contrast, type 3 verbalizations necessitate a search in the long-term memory and an elaboration about the motivations for a certain solutions path (Leighton, 2017, p. 41). In think-aloud interviews, type 1 and 2 verbalizations are typically conducted concurrently with the problem-solving processes. All the studies reviewed conducted concurrent think-aloud interviews during CT-related activities. However, it does not mean that only type 1 and 2 verbalizations have been involved. For example, one study uses think-aloud interviews to elicit learners' reflections scaffolding their CT problem-solving (Richard et al., 2018). Strictly speaking, the reflections are associated with type 3 verbalizations and are out of the scope of think-aloud interviews. In this case, cognitive laboratory interviews, instead of think-aloud interviews, should be used to collect the verbal data. Leighton (2017, p. 16) systematically distinguished between think-aloud interviews and cognitive laboratory interviews due to their distinctive objectives. While think-aloud interviews measure the problem-solving processes, cognitive laboratory interviews explore the understanding and comprehension processes. Cognitive laboratory interviews, which can be conducted either concurrently or retrospectively, are generally more flexible than think-aloud interviews (Leighton, 2017, p. 16). In fact, for the CT studies which aim to show participants' mastery of CT skills through verbalization, a cognitive laboratory interview would be a better option.

Although the two types of verbal data collection methods seem to be identical in surface aspects, they are different in objectives and procedures. The choice of the research method is not only related to the research purpose, but also associated with the properties of the target construct. Both think-aloud interviews and cognitive laboratory interviews are qualitative methods that provide verbal data. In this sense, both methods have great potential to present validity evidence for existing quantitative data and provide insights into participants' cognitive processes. Nevertheless, due to the different types of verbalizations elicited by the two methods, the choice of which method to use requires further reference to the specific purpose of the study, that is, whether the researchers were more interested in the comprehension or application of CT. Generally, it is difficult for one to apply a concept without comprehending it. However, the construct of CT has some special properties. On the one hand, the intensive contact with information technology in everyday life makes it possible that participants may apply CT unconsciously. For example, a participant may instinctively recognize the similarities and differences between a complex new problem and an old problem she/he has encountered before, even though she/he has no idea about the concept of pattern recognition, which is a component of CT. On the other hand, the construct of CT lacks a clear and unified definition. Tang et al. (2020) classified the operational definitions of CT into two main categories: (1) a set of skills related to programming and computing concepts; and (2) a competency that requires both domain-specific knowledge and general problem-solving strategies. Based on this classification, if researchers tend to define CT as a construct related to computing science or programming, it would be valuable to investigate participants' understanding and comprehension of CT concepts and practices. Cognitive laboratory interviews can provide rich data for this line of research. In contrast, if researchers tend to define CT as a problem-solving competency, it would be more valuable to study participants' cognitive processes in problem-solving, that is, the application of CT. Think-aloud interviews have great potential for examining the problem-solving processes, especially in a field that lacks clear models of cognition. For example, think-aloud interviews may allow researchers to discover the different thinking ways and strategies used by participants who are competent in CT and who are not. Think-aloud interviews play a pivotal role in research on expertise (Boshuizen & Schmidt, 1992; de Groot, 1965; Sabers et al., 1991). In the field of CT, think-aloud interviews may enable researchers to gain a better understanding of CT problem-solving at a cognitive level and thus design more effective interventions. Think-aloud interviews may also allow researchers to examine the baseline CT levels in participants with different backgrounds and domain knowledge at a greater depth. Although CT is believed to be closely related to several domain-specific skills, such as mathematical thinking and systematic thinking (Borkulo et al., 2020; Shute et al., 2017), there is still a lack of empirical evidence on how these abilities are linked at a cognitive level. Furthermore, when participants with limited prior knowledge of CT are solving problems that require the application of CT, they may transfer some of their domain knowledge and skills. Likewise, participants who have completed CT interventions may be asked to apply CT to solve problems in other domains (Ye et al., 2022). The above two situations represented the “transfer in” and “transfer out” processes proposed by Schwartz and Martin (2004). Understanding these two processes seems crucial to exploring the nature of human learning as well as developing lifelong learners in the fast-changing world. Think-aloud interviews may not only allow researchers to observe the dynamic process of CT knowledge transfer, but also provide inspiration regarding the individual differences (e.g., gender, age, culture) in CT. It is essential for researchers to review their objectives before collecting the verbal data and choose the most appropriate method according to the actual need.

Although most selected CT studies aim to investigate participants’ cognitive processes through think-aloud interviews, they tend to have different foci, leading to the divergence of analysis plans and outputs. For example, some studies use think-aloud interviews to identify the specific CT concepts and practices that participants had referred to during the problem-solving processes. In think-aloud sessions, participants’ CT proficiency was dependent on their tendency of mentioning certain concepts related to CS and programming, such as sequences, loops, and conditionals. Typically, the higher the frequency of CT-related concepts mentioned in think-aloud interviews, the higher the estimated level of participants’ CT proficiency. Participants’ CT proficiency is also assessed by knowledge and skill tests. In this scenario, think-aloud interviews are used to ensure the validity of traditional tests. Briefly speaking, to achieve a certain level of CT proficiency, participants not only have to complete the programming task, but they also need to show that they are attending to the corresponding concepts through verbalization. Some studies investigate which CT skills are present (e.g., Febrian et al., 2018; Lee et al., 2014), whereas others aim to see how participants adopt and articulate CT skills in various learning contexts, such as online debugging activities and educational games (e.g., Horn et al., 2016; Rowe et al., 2021; Vourletsis et al., 2021). These studies typically combine participants' verbal reports with observers' written notes during the analysis. Additionally, they pay attention to the integration of CT and domain knowledge. Another type of studies use think-aloud interviews to identify participants' challenges, misconceptions, and confusions in CT-related activities (e.g., Basu et al., 2016; Grover & Basu, 2017; Snow et al., 2017). Compared with traditional tests and questionnaire surveys, think-aloud interviews are more dynamic and malleable in nature (Cotton & Gresty, 2006; Merchie & Van Keer, 2014). As a consequence, they would have great potential to identify the transition and challenges in participants' learning trajectories.

Even though the implementing procedures of think-aloud interviews are similar across different regions and grade levels, the data analysis methods vary and lack standard guidelines. For instance, studies that adopted coding-and-counting strategies to analyze think-aloud data are based on different coding schemes (Basu et al., 2016; Jones-Harris & Chamblee, 2017; Lee et al., 2014). Despite the fact that self-developed coding schemes are likely to be more adaptable as they cater to the specific research context, the lack of a standard codebook may cause conflicts and misunderstandings, consequently impairing the validity of think-aloud interviews. Thus, further work is required to establish a standard analysis plan for CT-related think-aloud data. As suggested by Ericsson and Simon (1993), a cognitive model of problem-solving should be developed or identified before collecting the verbal data. Unfortunately, only a few CT studies have identified the cognitive models they use for coding in advance. The cognitive models for think-aloud interviews often involve some core dimensions of the broader CT conceptual framework, such as abstraction, debugging, and decomposition. There are also some attempts to develop the cognitive models based on the intersection of CT and other types of cognitive abilities, such as self-regulated learning and mathematical thinking. Given that think-aloud interviews are mainly used to uncover the cognitive processes embedded in CT, it is suggested that the cognitive perspective and time dimension of CT should be emphasized when developing the corresponding cognitive models. The think-aloud interview involves careful attention to specifying the cognitive processes and the probes, questions, and contexts that are used to collect the verbal data (Leighton, 2021). To improve the utility of think-aloud data, Leighton (2021) proposed three criteria requirements for clarifying the construct measured: (1) the construct is clearly defined as controlled cognitive processes that occur in working memory; (2) the construct is available in working memory for verbal reporting; and (3) the construct is measured in conditions that do not influence the memory of cognitive processes reported by the individual. Through reviewing some existing think-aloud studies, Leighton (2021) argued that most test constructs in these studies were not sufficiently well defined for in-depth investigation of the cognitive processes. Although some aspects of CT are expected to be available in working memory, whether all the cognitive processes involved in it are registered in working memory and verbalized is still unclear. Moreover, some think-aloud interviews in the reviewed studies were conducted in a group format, which could influence the cognitive processes measured during the CT-related task. As Leighton (2017, p. 12) suggested, the collaborative or communal aspect of the interview may bias the verbal responses provided by any one participant, which makes the claims about individual cognitive processes unreliable. To sum up, the definitional and methodological rigor should be established for future CT-related think-aloud studies, with the aim of making reasonable inferences and generating insightful information about participants' cognitive processes.

Chapter Summary

Chapter 6 summarizes the major findings of this study, identifying the trends and contexts of using think-aloud interviews. This chapter also notes the differences between various types of verbal data collection methods, highlighting the importance of choosing the appropriate protocols based on the research objectives. The chapter ends with a call for the establishment of standard guidelines for analyzing think-aloud data in CT studies. Overall, this chapter discusses the research findings from the review in the context of contemporary CT education and studies, providing a systematic overview of the application of think-aloud interviews.

Chapter 7 Educational Implications and Conclusion

This review contributes to existing knowledge of CT from both a theoretical and a practical perspective.

Theoretical Implications

From a theoretical perspective, it fills the gaps in the existing literature about CT assessment. Although think-aloud interviews have always been recommended to CT studies as assessment tools to unfold the cognitive processes, there is little systematic understanding of their theoretical foundations and implementing details. This review adopts a systematic framework to investigate the use of think-aloud interviews in CT studies. Through analyzing the characteristics of the related CT studies and exploring the incorporation of think-aloud interviews, it contributes to the knowledge about (1) the CT contexts in which think-aloud interviews can be conducted; (2) the roles of think-aloud interviews in CT studies; and (3) the analysis on the verbal data collected during CT-related activities. Additionally, this review identifies some limitations in the current CT studies using think-aloud interviews. The procedures in the selected articles depart to varying degrees from the guiding procedures of think-aloud interviews (Ericsson & Simon, 1993; Leighton, 2017). It is suggested that CT researchers should clearly identify a cognitive model before conducting think-aloud interviews. Moreover, CT researchers should employ more rigorous data collection and analysis protocols that conform to the accepted practices from decades of think-aloud research. Regarding the CT contexts, most think-aloud interviews were conducted in CS classrooms and involved only K-12 participants. In consideration of the significance of CT education and the importance of think-aloud interviews, more diverse educational settings and participants need to be considered in future studies. As an interdisciplinary concept in the information age, CT should be considered through various lenses, and it is necessary and important for researchers to synthesize knowledge from different domains.

Practical Implications

From a practical perspective, this review reveals that think-aloud interviews can provide researchers with new perspectives to study CT. Using think-aloud interviews, researchers can have a better observation of participants' real-time thoughts during CT problem-solving. They can also make inferences about participants' challenges, strategies, and individual differences in CT problem-solving from the think-aloud data (e.g., Basu et al., 2016; Horn et al., 2016; Jayathirtha et al., 2020). Consequently, the findings from think-aloud interviews may enable researchers to gain a deeper understanding of participants' cognitive processes and help them design more effective CT interventions. Think-aloud interviews used to be difficult to conduct in terms of time and scale. However, with the help of technology such as audio recording and screencasting, it is easier to collect and analyze think-aloud data, mitigating the obstacle of implementing think-aloud interviews to some extent (Mueller et al., 2017). With digital devices, think-aloud interviews can be conducted remotely. In particular, when participants are involved in the digital-based CT problem-solving, their behavioral and verbal data can be collected at the same time. Using online communication systems and screencasting software, educational institutions can collect the verbal data from a wider range of participants and analyze the data at any time. The large-scale analysis of verbal data has the potential to reveal the regional and longitudinal patterns of students' CT skills and attitudes. For educational institutions, this type of analysis may help them know better about students' CT competency over time and direct them to renew the current curriculum (Johnstone et al., 2006). For CT researchers, think-aloud interviews can serve as useful methods to understand how CT unfolds at the cognitive level, especially when combined with complementary assessment tools.

Future Research

The current review provides guidance for implementing think-aloud interviews in CT research. Importantly, it serves as a reference for studying the cognitive processes during CT problem-solving. The findings of the present review recommend researchers who intend to collect verbal data in CT studies to: (1) design CT activities according to the expected cognitive load. Researchers are encouraged to initiate a pilot study and expert reviews before collecting the verbal data. In the design phase, researchers can conduct cognitive laboratory interviews with a group of students to check on whether the CT activities are on the appropriate difficulty level and whether the language they used is understandable to the population of interest; (2) select appropriate interview probes based on research objectives. For example, if the research objective is to make inferences about participants' CT problem-solving processes, the probes are supposed to elicit participants' type 1 and 2 verbalizations, which involve the contents of participants' problem-solving in their working memory. In contrast, if the research objective is to explore participants' motivations for a given solution path in CT activities, the probes should elicit the type 3 verbalizations, which require participants to search their long-term memory, and then elaborate, explain, and clarify their response (Leighton, 2017, p. 41-42). Table 8 (Leighton, 2017, p. 44-45) presents an example of interview probes that can be used to elicit type 1 and 2 verbalizations during think-aloud interviews; and (3) identify or develop a cognitive model prior to data collection. Although CT is a field that lacks accepted models of cognition, there are some frameworks highlighting the cognitive/problem-solving processes in CT (e.g., Barr et al., 2011; Selby & Woollard, 2013). Researchers can refer to these frameworks and develop a cognitive model according to the specific problem-solving scenario. Leighton et al. (2004) developed a cognitive model of categorical reasoning based on the theory of mental models (Johnson-Laird, 1983). According to Leighton (2017, p. 103-107), this model not only guided the selection of tasks and participants for the following think-aloud interviews, but also guided the analysis of the collected verbal reports. In CT studies, for example, researchers can develop a cognitive model that involves a series of abilities ranging from abstracting to evaluating. Through the pilot study, they can identify some examples of utterances related to each ability in participants' verbal reports. Researchers can then decide on (1) the numerical or code values associated with each ability; (2) the rating levels representing the match between the ability and the utterance; and (3) the score levels for the accuracy of the utterance according to their cognitive models and research objectives (Leighton, 2017, p. 110-113).

Table 8

Example of Interview Probes in Think-Aloud Interviews

Category of Information/Purpose	Example Language in Instructions
1. Introduction to the objective of the study and the think-aloud interview procedure	In this study, we are interested in investigating the thoughts you have as you solve questions – specifically, computational thinking-related problems. For this reason, I am going to ask you to think aloud as you work through the five problems shown here.
2. Explanations of the think-aloud process – including parameters about not planning what is said or editing content	Let me explain what I mean by "think aloud". It means that I would like you to tell me everything you think about as you work through the computational thinking-related problems–we will do this one problem at a time. When I say tell me everything, I really mean every thought you have from the moment you read the problem to the end when you have a solution or even you do not have a solution. Please do not worry about planning how to say things or clarifying your thoughts–what I really what is to hear your thoughts constantly as you try to solve the problem–uninterrupted and unedited. Sometimes you may need time to think quietly through something–if so, that's ok but please tell me what you thought through as soon as possible after you are finished.
3. Creation of interview or environmental safety by acknowledging limitations of the procedure and explicitly indicating non-evaluation	I realize it can feel awkward to think aloud, but try to imagine you are alone in the room. If you become silent for too long, I will say "keep talking" as a way to remind you to think aloud. Also, please know that I am not an expert in this area so I cannot and will not be evaluating your thinking–the purpose of the study is to learn about the thoughts you have as you answer questions. We will have an opportunity to practice, but before we get to that, please let me know if you understand what we will be doing today? Do you have any questions? [Respond to any questions]
4. Implementation of practice problems	Let us now practice thinking aloud with some practice problems: Practice problem 1 [present to participant] Now, please tell me everything that you are thinking as you try to solve this.

Note. Adapted from *Using Think-Aloud Interviews and Cognitive Labs in Educational Research*, by J. P. Leighton, 2017, p. 44-45, Oxford University Press.

Furthermore, think-aloud interviews have the potential to identify competency gaps by involving diverse participants. For visually impaired participants, think alouds may be the most comfortable way to demonstrate their cognitive processes and CT proficiency. Additionally, CT skills and competencies are shown to be influenced by a series of factors, such as gender and age (Espino & González, 2015; Jiang & Wong, 2021; Witherspoon et al., 2016). For researchers who care about educational equity, it is imperative to investigate the underlying mechanism of distinctive CT patterns. As dynamic assessment tools, think-aloud interviews may be able to uncover the cognitive processes during CT problem-solving, providing the possibility for creating personalized CT profiles.

Conclusion

This systematic review examines the characteristics of CT-related think-aloud studies and explores the ways in which think-aloud interviews were incorporated into CT studies. The findings indicate that think-aloud interviews in CT-related studies are typically exploratory, conducted in CS classrooms with K-12 students, and are usually combined with other assessment tools. They provide convergent evidence about participants' CT competency and contribute to uncovering the cognitive processes embedded in CT. Moreover, think-aloud interviews have the potential to benefit learners with special needs and identify the competency gaps by involving diverse participants. This review suggests that, in the current CT research, more efforts need to be made to establish the definitional rigor of CT, to develop cognitive models, and to standardize the data analysis procedures for think-aloud interviews.

Chapter Summary

Chapter 7 discusses the educational implications of this study and provides guidelines for future CT research. Theoretically, this review highlights the limitations in the current CT studies using think-aloud interviews and suggests the need for researchers to clearly identify a cognitive model before conducting think-aloud interviews. Practically, this review serves as a reference for studying the cognitive processes involved in CT problem-solving and offers recommendations for CT researchers who intend to incorporate think-aloud interviews into their studies. This chapter ends with the conclusion of this systematic review.

References

Aho, A. V. (2012). Computation and computational thinking. *The Computer Journal*, *55*(7), 832–835. https://doi.org/10.1093/comjnl/bxs074

Aiken, J. M., Caballero, M. D., Douglas, S. S., Burk, J. B., Scanlon, E. M., Thoms, B. D., & Schatz, M. F. (2013). Understanding student computational thinking with computational modeling. *AIP Conference Proceedings*, *1513*(1), 46–49. https://doi.org/10.1063/1.4789648

Allan, V., Barr, V., Brylow, D., & Hambrusch, S. (2010). Computational thinking in high school courses. *Proceedings of the 41st ACM Technical Symposium on Computer Science Education - SIGCSE '10*, 390. https://doi.org/10.1145/1734263.1734395

Ambrosio, A. P., Xavier, C., & Georges, F. (2014). Digital ink for cognitive assessment of computational thinking. *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, 1–7. https://doi.org/10.1109/FIE.2014.7044237

Anderson, N. D. (2016). A call for computational thinking in undergraduate psychology. *Psychology Learning & Teaching*, *15*(3), 226–234. https://doi.org/10.1177/1475725716659252

Atmatzidou, S., & Demetriadis, S. (2014). How to support students' computational thinking skills in educational robotics activities. *Proceedings of 4th International Workshop Teaching Robotics, Teaching with Robotics & 5th International Conference Robotics in Education.* Padova, Italy.

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

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: What is Involved and what is the role of the computer science education community? *ACM Inroads*, *2*(1), 48–54. https://doi.org/10.1145/1929887.1929905

Basu, S., Biswas, G., Sengupta, P., Dickes, A., Kinnebrew, J. S., & Clark, D. (2016). Identifying middle school students' challenges in computational thinking-based science learning. *Research and Practice in Technology Enhanced Learning*, *11*(1). https://doi.org/10.1186/s41039-016-0036-2

Bell, T., Alexander, J., Freeman, I., & Grimley, M. (2009). Computer science unplugged: School students doing real computing without computers. *The New Zealand Journal of Applied Computing and Information Technology*, *13*(1), 20-29.

Bond, M. (2020). Facilitating student engagement through the flipped learning approach in K-12: A systematic review. *Computers & Education*, *151*, 103819. https://doi.org/10.1016/j.compedu.2020.103819

Bonner, S., Chen, P., Jones, K., & Milonovich, B. (2021). Formative assessment of computational thinking: cognitive and metacognitive processes. *Applied Measurement in Education*, *34*(1), 27–45. https://doi.org/10.1080/08957347.2020.1835912

Borkulo, S., Kallia, M., Drijvers, P., Barendsen, E., & Tolboom, J. (2020). Computational thinking and mathematical thinking: Digital literacy in mathematics curricula. *Proceedings of the 14th International Conference on Technology in Mathematics Teaching – ICTMT 14: Essen, Germany*, 384. https://doi.org/10.17185/DUEPUBLICO/70781

Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, *16*(2), 153–184. https://doi.org/10.1207/s15516709cog1602_1

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. *Proceedings of the 12th Workshop on Primary and Secondary Computing Education*, 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.* Paper. American Educational Research Association (AERA) Annual Meeting, Vancouver, BC, Canada.

Butler-Henderson, K., & Crawford, J. (2020). A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education*, *159*, 104024. https://doi.org/10.1016/j.compedu.2020.104024

Cotton, D., & Gresty, K. (2006). Reflecting on the think-aloud method for evaluating e-learning. *British Journal of Educational Technology*, *37*(1), 45–54. https://doi.org/10.1111/j.1467-8535.2005.00521.x

Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches* (3rd ed). Los Angeles, CA: Sage.

CSTA, & ISTE. (2011). *Computational Thinking in K–12 Education leadership toolkit*.

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

de Groot, A. (1965). *Thought and choice in chess* (2nd ed.). The Hague: Mouton Publishers.

del Olmo-Muñoz, J., Cózar-Gutiérrez, R., & González-Calero, J. A. (2020). Computational thinking through unplugged activities in early years of Primary Education. *Computers & Education*, *150*. https://doi.org/10.1016/j.compedu.2020.103832

Denner, J., Werner, L., Campe, S., & Ortiz, E. (2014). Using game mechanics to measure what students learn from programming games. *International Journal of Game-Based Learning*, *4*(3), 13–22. https://doi.org/10.4018/ijgbl.2014070102

Desoete, A. (2008). Multi-method assessment of metacognitive skills in elementary school children: How you test is what you get. *Metacognition and Learning*, *3*(3), 189–206. https://doi.org/10.1007/s11409-008-9026-0

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* The MIT Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. The MIT Press. https://doi.org/10.7551/mitpress/5657.001.0001

Espino, E. E. E., & González, C. S. G. (2015). Influence of Gender on Computational Thinking. *Proceedings of the XVI International Conference on Human Computer Interaction*, 1–2. https://doi.org/10.1145/2829875.2829904

Fan, M., Lin, J., Chung, C., & Truong, K. N. (2019). Concurrent think-aloud verbalizations and usability problems. *ACM Transactions on Computer-Human Interaction*, *26*(5), 1–35. https://doi.org/10.1145/3325281

Febrian, A., Lawanto, O., Peterson-Rucker, K., Melvin, A., & Guymon, S. E. (2018, June 24-27). Does everyone use computational thinking?: A case study of art and computer science majors. *ASEE Annual Conference & Exposition*, Salt Lake City, United States.

Fields, D. A., Lin, Y., Jayathirtha, G., & Kafai, Y. B. (2020). A Redesigned Reconstruction Kit for Rapid Collaborative Debugging and Designing of E-Textiles. *Proceedings of* *the FabLearn 2020 - 9th Annual Conference on Maker Education*, 98–101. https://doi.org/10.1145/3386201.3386207

Fisher, D., Frey, N., & Lapp, D. (2011). Coaching middle-level teachers to think aloud improves comprehension instruction and student reading achievement. *The Teacher Educator*, *46*(3), 231–243. https://doi.org/10.1080/08878730.2011.580043

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research*, *3*(4), 430–441. https://doi.org/10.1177/104973239300300403

Friend, M., & Cutler, R. (2013). Efficient egg drop contests: How middle school girls think about algorithmic efficiency. *Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research*, 99–106. https://doi.org/10.1145/2493394.2493413

Gardeli, A., & Vosinakis, S. (2017). Creating the computer player: An engaging and collaborative approach to introduce computational thinking by combining “unplugged” activities with visual programming. *Italian Journal of Educational Technology*, *1*(1). https://doi.org/10.17471/2499-4324/910

Gardeli, A., & Vosinakis, S. (2020). The effect of tangible augmented reality interfaces on teaching computational thinking: A preliminary study. In M. E. Auer & T. Tsiatsos (Eds.), *The Challenges of the Digital Transformation in Education* (Vol. 916, pp. 673–684). Springer International Publishing. https://doi.org/10.1007/978-3-030-11932-4_63

Grover, S., & Basu, S. (2017). Measuring student learning in introductory block-based programming: Examining misconceptions of loops, variables, and boolean logic. *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, 267–272. https://doi.org/10.1145/3017680.3017723

Guzdial, M. (2008). Education paving the way for computational thinking. *Communications of the ACM*, *51*(8), 25–27. https://doi.org/10.1145/1378704.1378713

Horn, B., Hoover, A. K., Barnes, J., Folajimi, Y., Smith, G., & Harteveld, C. (2016). Opening the black box of play: Strategy analysis of an educational game. *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, 142–153. https://doi.org/10.1145/2967934.2968109

Hsu, T.-C., Chang, S.-C., & Hung, Y.-T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. *Computers & Education*, *126*, 296–310. https://doi.org/10.1016/j.compedu.2018.07.004

Hu, J., & Gao, X. (2017). Using think-aloud protocol in self-regulated reading research. *Educational Research Review*, *22*, 181–193. https://doi.org/10.1016/j.edurev.2017.09.004

Huang, W., & Looi, C.-K. (2020). A critical review of literature on “unplugged” pedagogies in K-12 computer science and computational thinking education. *Computer Science Education*, 1–29. https://doi.org/10.1080/08993408.2020.1789411

India, G., Ramakrishna, G., Pal, J., & Swaminathan, M. (2020). Conceptual learning through accessible play: Project Torino and computational thinking for blind children in India. *Proceedings of the 2020 International Conference on Information and Communication Technologies and Development*, 1–11. https://doi.org/10.1145/3392561.3394634

Israel-Fishelson, R., & Hershkovitz, A. (2022). Studying interrelations of computational thinking and creativity: A scoping review (2011–2020). *Computers & Education*, *176*, 104353. https://doi.org/10.1016/j.compedu.2021.104353

Jayathirtha, G., Fields, D., & Kafai, Y. (2020). Pair debugging of electronic textiles projects: Analyzing think-aloud protocols for high school students’ strategies and practices while problem solving. In Gresalfi, M. and Horn, I. S. (Eds.), *The interdisciplinarity of the learning sciences, 14th International Conference of the Learning Sciences (ICLS) 2020, Volume 2* (pp. 1047-1054). Nashville, Tennessee: International Society of the Learning Sciences.

Jiang, S., & Wong, G. K. W. (2021). Exploring age and gender differences of computational thinkers in primary school: A developmental perspective. *Journal of Computer Assisted Learning*, jcal.12591. https://doi.org/10.1111/jcal.12591

Johnstone, C. J., Bottsford- Miller, N. A., & Thompson, S. J. (2006). *Using the think aloud method (cognitive labs) to evaluate test design for students with disabilities and English language learners* (Technical Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Johnson-Laird, P. N. (1983). *Mental models. Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Jones-Harris, C., & Chamblee, G. (2017). Understanding African-American students' problem-solving ability in the precalculus and advanced placement computer science classroom. In P. J. Rich & C. B. Hodges (Eds.), *Emerging Research, Practice, and Policy on Computational Thinking* (pp. 33–47). Springer International Publishing. https://doi.org/10.1007/978-3-319-52691-1_3

Kallia, M., van Borkulo, S. P., Drijvers, P., Barendsen, E., & Tolboom, J. (2021). Characterising computational thinking in mathematics education: A literature-informed Delphi study. *Research in Mathematics Education*. https://doi.org/10.1080/14794802.2020.1852104

Kazimoglu, C. (2020). Enhancing confidence in using computational thinking skills via playing a serious game: A case study to increase motivation in learning computer programming. *IEEE Access*, *8*, 221831–221851. https://doi.org/10.1109/ACCESS.2020.3043278

Kluge, A., Litherland, K. T., Borgen, P. H., & Lanslet, G. O. (2019). Combining programming with audio explanations. *Proceedings of the 2019 11th International Conference on Education Technology and Computers*, 155–159. https://doi.org/10.1145/3369255.3369283

Kruskopf, M., Lonka, K., & Mikkilä-Erdmann, M. (2019). Explorations on the nature of children's conceptual change in computational thinking. *INTED2019 Proceedings: 13th International Technology, Education and Development Conference*, 5884–5894. https://doi.org/10.21125/inted.2019.1441

Krutz, J., Siy, H., Dorn, B., & Morrison, B. B. (2019). Stepwise refinement in block-based programming. *Journal of Computing Sciences in Colleges*, 91–100.

K-12 Computer Science Framework Steering Committee. (2016). K-12 computer science framework. Technical Report. *Association for Computing Machinery*, New York, NY, USA.

Lajoie, S. P., Zheng, J., Li, S., Jarrell, A., & Gube, M. (2021). Examining the interplay of affect and self-regulation in the context of clinical reasoning. *Learning and Instruction*, *72*, 101219. https://doi.org/10.1016/j.learninstruc.2019.101219

Lee, T. Y., Mauriello, M. L., Ahn, J., & Bederson, B. B. (2014). CTArcade: Computational thinking with games in school age children. *International Journal of Child-Computer Interaction*, *2*(1), 26–33. https://doi.org/10.1016/j.ijcci.2014.06.003

Leighton, J., & Gierl, M. (Eds.). (2007). *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge University Press. https://doi.org/10.1017/CBO9780511611186

Leighton, J. P. (2017). *Using think-aloud interviews and cognitive labs in educational research*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199372904.001.0001

Leighton, J. P. (2021). Rethinking think-alouds: The often-problematic collection of response process data. *Applied Measurement in Education*, *34*(1), 61–74. https://doi.org/10.1080/08957347.2020.1835911

Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on tatsuoka's rule-space approach. *Journal of Educational Measurement*, *41*(3), 205–237. https://doi.org/10.1111/j.1745-3984.2004.tb01163.x

L'Heureux, J., Boisvert, D., Cohen, R., & Sanghera, K. (2012). IT problem solving: An implementation of computational thinking in information technology. *Proceedings of the 13th Annual Conference on Information Technology Education - SIGITE '12*, 183. https://doi.org/10.1145/2380552.2380606

Liu, R., Luo, F., & Israel, M. (2021). What do we know about assessing computational thinking? A new methodological perspective from the literature. *Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1*, 269–275. https://doi.org/10.1145/3430665.3456380

Liu, Y., & Rojas, J. (2019). Evaluation of the ROOT robot system and curriculum to improve computational thinking in Chinese children. *2019 IEEE R10 Humanitarian Technology Conference (R10-HTC)(47129)*, 126–131. https://doi.org/10.1109/R10-HTC47129.2019.9042435

Lockwood, J., & Mooney, A. (2018). Computational thinking in secondary education: Where does it fit? A systematic literary review. *International Journal of Computer Science Education in Schools*, *2*(1), 41–60. https://doi.org/10.21585/ijcses.v2i1.26

Looi, C.-K., How, M.-L., Longkai, W., Seow, P., & Liu, L. (2018). Analysis of linkages between an unplugged activity and the development of computational thinking. *Computer Science Education*, *28*(3), 255–279. https://doi.org/10.1080/08993408.2018.1533297

Luo, F., Israel, M., Liu, R., Yan, W., Gane, B., & Hampton, J. (2020). Understanding students' computational thinking through cognitive interviews: A learning trajectory-based analysis. *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 919–925. https://doi.org/10.1145/3328778.3366845

Lye, S. Y., & Koh, J. H. L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? *Computers in Human Behavior*, *41*, 51–61. https://doi.org/10.1016/j.chb.2014.09.012

Mardi, F. (2020). Using think alouds and digital powerups to embed computational thinking concepts while in-service teachers reflect on a math solution design project. *Journal of Digital Learning in Teacher Education*, *36*(4), 237–249. https://doi.org/10.1080/21532974.2020.1781001

McGee, S., McGee-Tekula, R., Duck, J., McGee, C., Dettori, L., Greenberg, R. I., Snow, E., Rutstein, D., Reed, D., Wilkerson, B., Yanek, D., Rasmussen, A. M., & Brylow, D. (2018). Equal outcomes 4 all: A study of student learning in ECS. *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, 50–55. https://doi.org/10.1145/3159450.3159529

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

Merchie, E., & Van Keer, H. (2014). Learning from text in late elementary education. Comparing think-aloud protocols with self-reports. *Procedia - Social and Behavioral Sciences*, *112*, 489–496. https://doi.org/10.1016/j.sbspro.2014.01.1193

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & for the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, *339*, b2535–b2535. https://doi.org/10.1136/bmj.b2535

Mueller, J., Beckett, D., Hennessey, E., & Shodiev, H. (2017). Assessing computational thinking across the curriculum. In P. J. Rich & C. B. Hodges (Eds.), *Emerging Research, Practice, and Policy on Computational Thinking* (pp. 251–267). Springer International Publishing. https://doi.org/10.1007/978-3-319-52691-1_16

Ness, M. (2016). Learning from K-5 teachers who think aloud. *Journal of Research in Childhood Education*, *30*(3), 282–292. https://doi.org/10.1080/02568543.2016.1178671

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Prentice-Hall.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Papert, S. (1980). Mindstorms: children, computers, and powerful ideas. *Mindstorms,* New York: Basic Books.

Papert, S. (1991). Situating constructionism. In S. Papert, & I. Harel (Eds.), *Constructionism.* Cambridge, MA: MIT Press.

Pasternak, A. (2016). Contextualized teaching in the lower secondary education long-term evaluation of a CS course from grade 6 to 10. *Proceedings of the 47th ACM Technical Symposium on Computing Science Education - SIGCSE '16*, 657–662. https://doi.org/10.1145/2839509.2844592

Richard, G. T., Giri, S., McKinley, Z., & Ashley, R. W. (2018). Blended making: Multi-interface designs and e-crafting with elementary and middle school youth. *Proceedings of the 17th ACM Conference on Interaction Design and Children*, 675–680. https://doi.org/10.1145/3202185.3210798

Rodriguez, B., Kennicutt, S., Rader, C., & Camp, T. (2017). Assessing computational thinking in CS unplugged activities. *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, 501–506. https://doi.org/10.1145/3017680.3017779

Rowe, E., Almeda, M. V., Asbell-Clarke, J., Scruggs, R., Baker, R., Bardar, E., & Gasca, S. (2021). Assessing implicit computational thinking in Zoombinis puzzle gameplay. *Computers in Human Behavior*, *120*. https://doi.org/10.1016/j.chb.2021.106707

Sabers, D. S., Cushing, K. S., & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensional, and immediacy. *American Educational Research Journal*, *28*(1), 63–88. https://doi.org/10.3102/00028312028001063

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, *22*(2), 129–184. https://doi.org/10.1207/s1532690xci2202_1

Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 73–163). The Guilford Press.

Selby, C., & Woollard, J. (2013). Computational thinking: The developing definition. Retrieved from https://people.cs.vt.edu/~kafura/CS6604/Papers/CT-Developing-Definition.pdf

Sengupta, P., Kinnebrew, J. S., Basu, S., Biswas, G., & Clark, D. (2013). Integrating computational thinking with K-12 science education using agent-based computation: A theoretical framework. *Education and Information Technologies*, *18*(2), 351–380. https://doi.org/10.1007/s10639-012-9240-x

Seo, J., & Richard, G. T. (2018). Accessibility, making and tactile robotics: Facilitating collaborative learning and computational thinking for learners with visual impairments. *Proceedings of International Conference of the Learning Sciences, 3*, 1755–1756.

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

Snow, E. B., Rutstein, D. W., Bienkowskie, M., & Xu, Y. (2017). Principled assessment of student learning in high school computer science. *Proceedings of the 2017 ACM Conference on International Computing Education Research* (pp. 209-216). https://doi.org/10.1145/3105726.3106186

Solomon, P. (1995). The think aloud method: A practical guide to modelling cognitive processes. *Information Processing & Management*, *31*(6), 906–907. https://doi.org/10.1016/0306-4573(95)90031-4

Starbird, K., & Palen, L. (2011). More than the usual suspects: The physical self and other resources for learning to program using a 3D Avatar environment. *Proceedings of the 2011 iConference*. 614–621. https://doi.org/10.1145/1940761.1940844

Statter, D., & Armoni, M. (2020). Teaching abstraction in computer science to 7th grade students. *ACM Transactions on Computing Education*, *20*(1), 8–837. https://doi.org/10.1145/3372143

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

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 the 14th Workshop in Primary and Secondary Computing Education*, 1–9. https://doi.org/10.1145/3361721.3361729

Turchi, T., Fogli, D., & Malizia, A. (2019). Fostering computational thinking through collaborative game-based learning. *Multimedia Tools and Applications*, *78*(10), 13649–13673. https://doi.org/10.1007/s11042-019-7229-9

Vourletsis, I., Politis, P., & Karasavvidis, I. (2021). The effect of a computational thinking instructional intervention on students' debugging proficiency level and strategy use. In T. Tsiatsos, S. Demetriadis, A. Mikropoulos, & V. Dagdilelis (Eds.), *Research on E-Learning and ICT in Education* (pp. 15–34). Springer International Publishing. https://doi.org/10.1007/978-3-030-64363-8_2

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

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

Witherspoon, E. B., Schunn, C. D., Higashi, R. M., & Baehr, E. C. (2016). Gender, interest, and prior experience shape opportunities to learn programming in robotics competitions. *International Journal of STEM Education*, *3*(1). https://doi.org/10.1186/s40594-016-0052-1

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). https://doi.org/10.1145/2576872

Yadav, A., Ocak, C., & Oliver, A. (2022). Computational thinking and metacognition. *TechTrends*, *66*(3), 405–411. https://doi.org/10.1007/s11528-022-00695-z

Ye, J., Lai, X., & Wong, G. K. (2022). The transfer effects of computational thinking: A systematic review with meta-analysis and qualitative synthesis. *Journal of Computer Assisted Learning*, jcal.12723. https://doi.org/10.1111/jcal.12723

Yuen, T. T., & Robbins, K. A. (2014). A qualitative study of students' computational thinking skills in a data-driven computing class. *ACM Transactions on Computing Education*, *14*(4). https://doi.org/10.1145/2676660