ABSTRACT

RACHMATULLAH, ARIF. Learning and Teaching about Food Webs in a Computationally Rich Environment: A Mixed-Methods Study. (Under the direction of Dr. Eric N. Wiebe).

This dissertation conducted a science classroom intervention using two instructional approaches, computational modeling and paper-based pictorial modeling, in the context of food webs. A series of research papers were written on the impact of the intervention on students' attitudes and learning, and on teachers via professional development and teaching. The first study employed quasi-experimental design, examining the impacts of computational modeling and paper-based pictorial modeling on students' systems thinking and CT skills. A total of 365 seventh-grade students were involved in week-long online learning activities. The students were purposefully assigned to either synchronous computational modeling condition (n = 224) or asynchronous paper-based pictorial modeling condition (n = 141). Students modeled the interrelation between concepts and components of a food web in both conditions. The students took CT and systems thinking-embedded food web assessments before and after the activity and formative assessment after each activity to gather their perceptions of what they thought they learned from the activities. Multilevel modeling and Epistemic Network Analysis were used to analyze the data. The findings indicated a significant increase in students' systems thinking skills regardless of the conditions. Students in computational modeling condition had a substantial increase in their CT scores, while such a result was not detected in the paper-based pictorial modeling condition. ENA results supported this finding, showing that students in computational modeling condition expressed that they learned both science and CT concepts.

The second study explored the predictors of students' interests in computationally intensive (or AI-informed) science careers and the impact of computational modeling activity on middle school students' interests in such careers. In parallel, an instrument to measure students'

career interests called the Computationally Intensive Science Career Interest (CISCI), was developed and validated. A combination of classical test theory and item response theory was used to validate the CISCI instrument using a sample $N = 934$. Multiple linear regression and paired-sample *t*-tests were performed to analyze the data. The results showed that the CISCI is a valid and reliable instrument consisting of five constructs. Results from regression tests revealed that students' science attitudes, computer science attitudes, CT skills, and prior computer science-related activities significantly predicted students' career interests. Furthermore, a significant increase in students' perceptions of intensity of discussing computationally intensive careers with their parents was detected after participating in a computational modeling activity. The third study employed a mixed-methods design to investigate the changes in, and sources of, middle school teachers' self-efficacy for teaching science in a computationally rich environment. A total of eleven middle school teachers participated in this study. They took two questionnaires four times—before and after training and teaching—measuring their self-efficacy for teaching science and CT. The teachers were also either interviewed or asked to provide written reflections after taking the questionnaires. Non-parametric Skillings-Mack tests and thematic analysis were used to analyze the data. The results revealed a significant increase in teachers' self-efficacy for teaching science and CT after training and teaching. These changes in self-efficacy were tied to three sources, namely: (1) exposure to computer programming, (2) students' interests and responses to a computationally rich science environment, and (3) teaching repetition and field experience. Results from the three studies highlight the potential benefit of engaging students with learning science in a computationally rich environment to positively influence both their CT and systems thinking skills as well as content knowledge. The study also highlighted the

challenges and possibilities for preparing science teachers to teach computational modeling with little or no background in programming or computer science.

Learning and Teaching about Food Webs in a Computationally Rich Environment: A Mixed-Methods Study

by
Arif Rachmatullah

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Learning and Teaching in STEM

Raleigh, North Carolina
2021

APPROVED BY:

Dr. Eric N. Wiebe Committee Chair	Dr. M. Gail Jones
Dr. Soonhye Park	Dr. Christopher B. Mayhorn

DEDICATION

This dissertation is dedicated to all the incredible teachers and students who are enthusiastic about innovative instructional approaches yet still lack access to such resources.

BIOGRAPHY

Arif was born and raised in Sumedang, Indonesia. He then moved to the neighboring city, Bandung, to earn a biology education degree from the Universitas Pendidikan Indonesia. His research interest grew tremendously when writing his undergraduate thesis on the scientific literacy of middle school students in his hometown. Since then, he has committed himself to research and moved to Korea to build his science education research experience, where Dr. Minsu Ha advised him. Arif was fortunate to have Dr. Ha as his advisor, who supported Arif's every move to develop his identity as a researcher by involving him in numerous research projects. These opportunities allowed Arif to publish more than five journal articles by his master's graduation. Upon completing his master's degree at Kangwon National University, Arif continued studying in the STEM education program at North Carolina State University in Raleigh, NC. He refined his research identity while working on his doctoral study, advised by Dr. Eric Wiebe. Arif believes that his greatest accomplishment is contributing to the STEM Cyberlearning Team. Afterward, he finally returned to classrooms and observed students and teachers learning and teaching science in a computationally rich environment. Arif is dedicated to contributing to this emerging research area and committed as a researcher at the intersection of science education, computer science education, and educational technology. Arif looks forward to leveraging the research knowledge and skills he gains throughout his studies to contribute to this research area and apply them to his next journey.

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to all remarkable people who, without them, this dissertation work could not have happened. First and foremost, I would like to thank my incredible advisor and dissertation chair, Dr. Eric Wiebe, for all encouragement, constructive feedback, brainstorming, and challenges that helped me grow as a person and researcher. I am beyond happy to have the opportunity to work with you that made my three years at NC State enjoyable and fantastic. Thank you for always listening to all my thoughts and supporting all my decisions. I also want to thank my committee members, Dr. M. Gail Jones, Dr. Soonhye Park, and Dr. Christopher Mayhorn, who always give positive energy, encouragement, and constructive feedback to my work.

Second, I would like to thank all participating students, parents, and teachers. Without their time and willingness, I would not be able to have this quality of dissertation work. I am so thankful for all 1611 consented and not consented students and eleven teachers who participated in the study for bearing with a new instructional approach amid difficult pandemic time. Thank you, my eleven teachers, for constantly reminding students to finish their work and ask parents to sign the consent forms. Thank you, students' parents, for letting your children being part of my study.

Next, I should thank all my STEM Cyberlearning team members for all their supports and laughs. All of you always make my day. You all deserve to be mentioned: Dolly Bounajim, Madeline Hinckle, Zarifa Zakaria, Dr. Jessica Vandenberg, Dr. Danielle Boulden, Dr. Bita Akram, Dr. Cody Smith, and Alonzo Alexander. I also must thank all people over the Center for Educational Informatics (CEI) led by Dr. James Lester and University of Florida's

LearnDialogue Group led by Dr. Kristy Boyer. I am also grateful to all people at the Friday Institute for Educational Innovation.

I would also like to acknowledge all my science education cohort for their support and positive energy throughout my doctoral journey. Thanks to all my Indonesian friends at NC State for all the excellent food, games, and laughs. I should admit that I could have gone this far because of all of you. To Ai and Yustika, thanks for being my second coders, coding my qualitative data overnights.

Finally, I must thank my family. Thank you, Mamah and Ayah, for always supporting all my journey. Thank you for constantly worrying me, asking the same questions whether I have eaten something or not. Thank you for asking me whether I am healthy or not. Those repeated, annoying and straightforward questions were actually something special to me that kept me from homesick. Thank you to my sisters and brother for keeping me updated with all news and gossip in my home country.

Special acknowledgment to NC State College of Education and Friday Institute for awarding me the funding to conduct this study through the Dissertation Research Continuity program and Graduate Student Fellows.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Introduction

The growing number of computationally intensive science careers has driven many changes in science education instruction. Such changes have focused on fostering cognitive, affective-motivational, and skills in K–12 science education aiming for preparing students' readiness to enter such careers. Skills and abilities such as computational thinking (CT) and systems thinking have been included as core components of science curricula in many countries. Such thinking skills and disciplinary knowledge are thought to be the essential skills in computationally intensive science careers. Emerging research has been exploring different instructional approaches that enable the advancement of such thinking skills. These policy shifts have also pointed to the need to help teachers prepare to use such instructional approaches to help students develop those skills. This dissertation attempted to explore how middle school students and teachers learn and teach in a computationally rich environment.

Statement of the Problem

With the increasing reliance on technology and computation, the way that society reasons through and solves problems has changed in the past few decades. Research is now suggesting that students develop and nurture computational thinking (CT) skills—defined as a set of skills related to thinking abstractly and algorithmically—using computer science logic, concepts, and skills to solve problems, though not necessarily involving computers (Shute et al., 2017; Wing, 2006). The Next Generation Science Standards (NGSS) considers CT one of eight essential science and engineering practices in which students should be engaged throughout their education (NGSS Lead States, 2013). Building literacy in CT is regarded as essential for students to excel in, and pursue advanced science, technology, engineering, and mathematics (STEM)

degrees and careers (Leonard et al., 2016). Such computational skills and techniques were central to the record-breaking speed at which the COVID-19 vaccines were developed (Arnold, 2020).

The incorporation of CT into STEM curricula, particularly science, can be achieved through the implementation of computational modeling activities in which students work to build scientific models through computer programming (e.g., Aksit & Wiebe, 2020; Irgens et al., 2020). When students are involved in these activities, they learn to view scientific problems holistically through an algorithmic lens, facilitating the generation of many possible solution scenarios (Peel et al., 2019; Weintrop et al., 2016). As an example, when students solve problems related to the carbon cycle, they identify all components in the carbon cycle, generate hypotheses and underlying reasoning, and iteratively test their hypotheses by building computational models. Moreover, these activities encourage students to use computer science concepts, such as logical conditionals, that are also an essential part of (hypothetico-deductive) scientific reasoning (Lawson, 2010). The emergence of the research and policy focus on this practice has triggered a debate around how CT differs from other thinking skills, particularly scientific reasoning and systems thinking (Shute et al., 2017).

While some studies have demonstrated that CT is intertwined with scientific reasoning, as both skill sets foster conditional, abstract, and algorithmic thinking (e.g., Peel et al., 2019), other scholars have theoretically conceptualized CT as a broader construct than the aforementioned scientific reasoning skills (Shute et al., 2017; Weintrop et al., 2016). There is still minimal research exploring how CT and scientific reasoning relate to one another. We believe that such studies are imperative, primarily to provide more evidence to support and guide the practice of CT in STEM curricula. Moreover, given that scientific reasoning sits at the foundation of scientific literacy and thus a primary objective of science education (Bybee &

Roberts, 2014; Fives et al., 2014), examining how CT relates to scientific reasoning could help achieve this objective.

A facet of scientific reasoning is systems thinking, defined as the ability to think in a systemic context by focusing on the interactions and relationships among the components of a system and reasoning across time and scales (Jin et al., 2019; Mambrey et al., 2020a; Senge, 1990). Systems thinking has emerged in the nascent CT research literature as a particularly synergistic point of intersection between scientific reasoning and CT (Weintrop et al., 2016). Systems thinking is often used when explaining large-scale scientific phenomena, such as ecosystems and climate systems. The use of systems thinking is pivotal in science, given that all scientific phenomena are built on a part of a more extensive system. Even though research on both CT and systems thinking is limited, some studies have noted that the two share certain common features (Shute et al., 2017; Weintrop et al., 2016), such as the concept of "loops" (Sweeney & Sterman, 2007). This proposed research is interested in better understanding the relationship between systems thinking and CT skills. More specifically, exploring this relationship in the context of ecosystems. For example, whether students with higher CT skills explain the concept of ecosystems by accounting for more components or viewing them as more interconnected and complex than those students with a lower level of CT skills (cf., Jin et al., 2019; Mambrey et al., 2020a). Furthermore, this study would examine how participation in scientific computational modeling activities influences these two constructs.

As mentioned earlier, the increased interest in CT in STEM education has promoted the development of computationally rich science activities, with the goal to enrich students' science learning experiences, particularly with computational models and modeling practices. One of the goals for exposing students to these activities is to support the development of students' interest

in computationally-intensive science careers (Falloon et al., 2020; Goode & Chapman, 2016), which are in increasingly high demand (Navlakha & Bar-Joseph, 2011). However, few CT-related studies have looked at the impact of such interventions on computationally-intensive careers, which Bortz et al. (2019) attributed, in part, to a lack of psychometrically sound instruments. Thus, a final goal of this dissertation work is the validation of an instrument that can be used to examine the impact of students' participation in computationally rich science activities on such career interests.

Not only does the inclusion of CT into science curricula influence students, teachers are also affected by the inclusion of computationally-intensive activities in classrooms. Many researchers have reported the need for teachers' professional development (PD) programs focusing on CT-integrated science teaching and learning (e.g., Ketelhut et al., 2020; Langbeheim et al., 2020). Such PD programs are necessary, in part for raising teachers' self-efficacy to teach science in a computationally rich environment. This is a focus in part because CT and computer science concepts are not part of a majority of teacher preparation programs (Langbeheim et al., 2020). This need is reinforced by research showing that teachers' teaching efficacy beliefs are the function of students' learning (Tschannen-Moran & Barr, 2004; Zee & Koomen, 2016).

The inclusion of CT/CS into science curricula has given rise to a new symbol/language, which is computer codes, to represent a scientific phenomenon, in the following section the learning theories used in this study are described.

Theoretical Framework: Symbolic Tools and Constructionism

For Vygotsky, learning cannot happen without mediations using tools and signs. Even though tools may include physical tools, Vygotsky also points to psychological tools such as mnemonic devices (Wertsch, 2009). Additionally, he was more interested in the relationships

between the role of signs (semiotics), including human language, and how they mediate human actions (Wertsch, 2009). Vygotsky believed that "it is meaningless to assert that individuals have a sign, or have mastered it, without addressing the ways in which they do or do not use it to mediate their own actions or those of others" (Wertsch, 2009, p. 29). This implies that Vygotsky's view of signs may be connected to the process of meaning-making. This aligns with Peirce's (1998) pragmatist view of semiosis (meaning-making) that explains the continuous process in which an individual generates, synthesizes and shares meaning through linking signs, objects, and real-world contexts. Furthermore, Vygotsky argues that certain languages may trigger different forms of thinking (Wertsch, 2009), which makes language unique with regard to mental processing.

From the perspective of constructionism, learning is designing and building objects that embody the structural ideas of a content domain. DiSessa (1979) argues that these objects can be in the form of "microworlds" that represent "the powerful logical structure of the subject, but which at the same time mesh properly with the cognitive reality of human beings" (p. 239). Papert (1980) asserts that students who are involved in constructionist learning activities typically use an authoring environment where they design and program artifacts that constitute the underlying structures and rules of a phenomenon. Papert adds that the learners usually articulate their prior knowledge in the form of computer codes and then actively revise them through debugging activities. This active revision of codes, a sign system, is equivalent to the revision of conceptions, which enables learners to generate refined knowledge or understanding. The introduction of microworlds, in the form of programmable computational models, introduces a new set of tools and signs through which students can mediate their meaning-making. Wilensky (2003), Kynigos (2007) and others have advanced constructionism by adding new

tools and techniques that allow students in a constructionist learning environment to interact with models by revising or adding rules, variables or entities to an existing program or pre-built model, which they refer to as “extensible models” or “half-baked microworlds.”

Research Purposes and Questions

This dissertation aims to examine: (1) the influence of computational modeling activities on students’ CT and systems thinking skills; (2) the development of an instrument to measure students’ interest in computationally demanding science careers, and the impact of students’ participation in computationally rich science activities on their interest in careers in these areas; and (3) the developmental trends and sources of teachers’ self-efficacy for teaching science and CT. These aims are addressed in three different studies, which are described in greater detail in Chapter 2, Chapter 3, and Chapter 4. The three studies in this dissertation are answering the following three research questions:

1. What are the impacts of computational modeling activities on students’ systems thinking and computational thinking skills?
2. To what extent does participating in computationally rich science activities influence students’ interests in computationally intensive science careers?
3. How do teachers’ self-efficacy for teaching science look like when they teach in a computationally rich environment for the first time?

Significance of the Research

The three studies in this dissertation can benefit curriculum developers, education researchers, and teachers. The first study presents an example of students’ learning in computationally rich science activities through 3D model of learning (disciplinary content knowledge, science and engineering practices, and crosscutting concepts). This study

demonstrates the connections among systems-thinking skills, conceptual understandings of ecosystems, and CT skills supported by rich empirical data. Moreover, this first study shows how computationally rich science activities can advance students' conceptual understanding of science and CT skills and aid students in viewing scientific phenomena from a systems perspective.

The second study validates a psychometrically sound instrument that educational researchers and teachers can use to evaluate the effectiveness of computationally rich science interventions in influencing career interest. In addition to its use by researchers, policymakers can also use data from this instrument to guide interventions designed to increase the number of students on computationally-intensive STEM career pathways. This instrument responds to our evolving digital era, where jobs in computational sciences are increasing in number, especially in STEM areas such as science. This study uncovers the predictors of student interest in these kinds of jobs, which can guide policymakers, educational researchers, and teachers toward interventions that target strategies that increase these important predictors.

Finally, the third study provides evidence for the importance of a PD program that introduces the basic concepts and skills of computer programming to teachers so that they are familiar with such concepts and skills. This study also shows how experience teaching science in a computationally rich environment can increase teachers' teaching efficacy beliefs. The third study underscores the need for more PD and teaching experience for teachers so that they are better prepared to teach in a computationally rich teaching environment.

REFERENCES

Aksit, O., & Wiebe, E. N. (2020). Exploring force and motion concepts in middle grades using computational modeling: A classroom intervention study. *Journal of Science Education and Technology*, *29*(1), 65–82. https://doi.org/10.1007/s10956-019-09800-z

Arnold, C. (July, 2020). How computational immunology changed the face of COVID-19 vaccine development. *Nature Medicine*. https://doi.org/10.1038/d41591-020-00027-9

Bortz, W. W., Gautam, A., Tatar, D., & Lipscomb, K. (2020). Missing in measurement: Why identifying learning in integrated domains is so hard. *Journal of Science Education and Technology*, *29*(1), 121–136.

DiSessa, A. (1979). On 'learnable' representations of knowledge: A meaning for the computational metaphor. In J. Lochhead & J. Clement (Eds.), *Cognitive process instruction: Research on teaching thinking skills* (pp. 239–266). Philadelphia, PA: The Franklin Institute Press.

Falloon, G., Hatzigianni, M., Bower, M., Forbes, A., & Stevenson, M. (2020). Understanding K-12 STEM education: a framework for developing STEM literacy. *Journal of Science Education and Technology*, *29*(3), 369–385. https://doi.org/10.1007/s10956-020-09823-x

Fives, H., Huebner, W., Birnbaum, A. S., & Nicolich, M. (2014). Developing a measure of scientific literacy for middle school students. *Science Education*, *98*(4), 549–580. https://doi.org/10.1002/sce.21115

Goode, J., & Chapman, G. (2016). *Exploring computer science*. University of Oregon, Eugene, OR. Retrieved September 11, 2020 from https://www.teach21.us/uploads/1/3/0/5/13053428/samplecscurriculum.pdf

Irgens, G. A, Dabholkar, S., Bain, C., Woods, P., Hall, K., Swanson, H., Horn, M., & Wilensky, U. (2020). Modeling and measuring high school students' computational thinking practices in science. *Journal of Science Education and Technology*, *29*(1), 137–161. https://doi.org/10.1007/s10956-020-09811-1

Jin, H., Shin, H. J., Hokayem, H., Qureshi, F., & Jenkins, T. (2019). Secondary students' understanding of ecosystems: A learning progression approach. *International Journal of Science and Mathematics Education*, *17*(2), 217–235. https://doi.org/10.1007/s10763-017-9864-9

Ketelhut, D. J., Mills, K., Hestness, E., Cabrera, L., Plane, J., & McGinnis, J. R. (2020). Teacher change following a professional development experience in integrating computational thinking into elementary science. *Journal of Science Education and Technology*, *29*(1), 174–188. https://doi.org/10.1007/s10956-019-09798-4

Kynigos, C. (2007). Using half-baked microworlds to challenge teacher educators' knowing. *International Journal of Computers for Mathematical Learning*, *12*(2), 87–111. https://doi.org/10.1007/ s10758-007-9114-2

Langbeheim, E., Perl, D., & Yerushalmi, E. (2020). Science Teachers' Attitudes towards Computational Modeling in the Context of an Inquiry-Based Learning Module. *Journal of Science Education and Technology*, *29*(6), 785–796. https://doi.org/10.1007/s10956-020-09855-3

Lawson, A. E. (2010). Basic inferences of scientific reasoning, argumentation, and discovery. *Science Education*, *94*(2), 336–364. https://doi.org/10.1002/sce.20357

Leonard, J., Buss, A., Gamboa, R., Mitchell, M., Fashola, O. S., Hubert, T., & Almughyirah, S. (2016). Using robotics and game design to enhance children's self-efficacy, stem attitudes,

and computational thinking skills. *Journal of Science Education and Technology*, *25*(6), 860–876. https://doi.org/10.1007/s10956-016-9628-2

Mambrey, S., Timm, J., Landskron, J. J., & Schmiemann, P. (2020). The impact of system specifics on systems thinking. *Journal of Research in Science Teaching*, 1–20. https://doi.org/10.1002/tea.21649

Navlakha, S., & Bar-Joseph, Z. (2011). Algorithms in nature: The convergence of systems biology and computational thinking. *Molecular Systems Biology*, *7*, 1-11. https://doi.org/10.1038/msb.2011.78

NGSS Lead States. (2013). *Next generation science standards. For states, by states*. Washington, DC: The National Academies Press.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York, NY, USA: Basic Books, Inc.

Peel, A., Sadler, T. D., & Friedrichsen, P. (2019). Learning natural selection through computational thinking: Unplugged design of algorithmic explanations. *Journal of Research in Science Teaching*, *56*(7), 983–1007. https://doi.org/10.1002/tea.21545

Peirce, C. S. (1998). What is a sign? In N. Houser, A. De Tienne, J. R. Eller, C. L. Clark, A. C. Lewis, & D. B. Davis (Eds.), *The essential Peirce—Selected philosophical writings* (1893–1913) (Vol. 2, pp. 4–10). Bloomington, IL: Indiana University Press.

Roberts, D. A., & Bybee, R. W. (2014). Scientific literacy, science literacy, and science education. In S. K. Abell,&N. G. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 545–558). New York, NY: Routledge.

Senge, P. M. (1994). *The fifth discipline: The art and practice of the learning organization*. New York, NY: Doubleday.

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

Tschannen-Moran, M., & Barr, M. (2004). Fostering student learning: The relationship of collective teacher efficacy and student achievement. *Leadership and Policy in Schools*, *3*(3), 189–209. https://doi.org/10.1080/15700760490503706

Vygotsky, LS. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

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

Wertsch, J. V. (2009). *Voices of the mind: Sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.

Wilensky, U. (2003). Statistical mechanics for secondary school: The GasLab modeling toolkit. *International Journal of Computers for Mathematical Learning*, *8*(1), 1–41.

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

Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015. https://doi.org/10.3102/0034654315626801

CHAPTER 2: Building a Computational Model of Food Webs: Impacts on Middle School Students' Computational and Systems Thinking Skills

Abstract

Integral to fostering computational thinking (CT) skills, which are increasingly essential in today's digital era, has been the shift of paper-based modeling activities to computational modeling. Research has indicated that modeling activities can advance students' understanding of a system's mechanism (i.e., systems thinking), such as an ecosystem. The current study examines the impacts of paper-based and computational modeling activities on students' systems thinking and CT skills. A total of 365 seventh-grade students were involved in online modeling activities, spanning over four days, and were assigned purposefully to a paper-based modeling (n = 141) or a computational modeling (n = 224) condition. They took systems-thinking-embedded food web and CT assessments before and after the four activities, in addition to a formative assessment after each activity. Multilevel modeling and repeated-measures correlation tests were used to analyze the students' quantitative data. Epistemic Network Analysis (ENA) was utilized to map out students' perceptions of what they believed they learned from the activities. The results revealed significant increases in two out of three systems thinking-embedded food web constructs (systems organization and behavior) in both conditions. However, the increase of CT skills in the paper-based pictorial/drawing condition was not as significant as the increase in the computational modeling condition. ENA results showed that students in the computational modeling condition had more co-occurrences between science and computer science or CT concepts than those in the paper-based condition. These findings illuminate the benefit of engaging students in computationally rich science activities to advance both systems thinking and CT skills.

Keywords: computational thinking, food web, middle school, systems thinking

Introduction

The objectives of science and science education include building models and modeling activities to understand particular scientific phenomena. Scientific models that students and scientists create can take diverse forms, including tables, diagrams, symbols, codes, photographs, simulations, and gestures (Gilbert, 2008; Suàrez, 2003). Research has suggested that fluency and flexibility of using different forms of model in a scientific explanation, called representational competence, show possession of deep conceptual understanding (Anderson et al., 2013; Gilbert & Treagust, 2009; Pande & Chandrasekharan, 2017). In addition, engaging students in any types of modeling activities have been found to support the development of advanced thinking skills, such as systems thinking (Nguyen & Santagata, 2020), which is among essential crosscutting concepts listed in the Next Generation Science Standards (NGSS Lead States, 2013).

Systems thinking is a critical skill for students to learn to identify all features and operations of the represented scientific phenomena and the interconnections between one element and another and across concepts (Mehren et al., 2018; Wilensky & Resnick, 1999). Identifying all the features of a system and how they are interconnected is an essential competence for learning about ecological concepts, such as food webs (Mambrey et al., 2020a). This is evident in several studies demonstrating that students were prone to recognizing the discrete phenomenon of predator and prey in a food web but could not see the other relationships (Eberbach et al., 2012; Eilam, 2012; Hogan, 2000; Reiner & Eilam, 2001). As a further example, students could identify the “eating order” within food webs—for instance, that wolves eat rabbits—but tended not to see the second-order impact on grasses’ population. Reiner and Eilam (2001) described this reasoning phenomena as incoherent but consistent, meaning that students

incoherently relate the process of feeding to other components of food webs, as well as other concepts (e.g., energy transmission), but are consistent in their application of the concept of eating order. Therefore, advancing students' systems thinking skills through modeling activities is critical to achieving advanced conceptual understanding about food webs.

Recently, modeling activities have been shifting from traditional modeling using physical models or drawing towards more computer-based activities (Coll & Lajium, 2011). Emergent research has now suggested engaging K-12 students in computationally rich science activities (e.g., when students work to build scientific models through computer programming) to help enrich students' abilities to create representations of scientific concepts (Moore et al., 2020) as well as systems thinking (Mahaffy et al., 2019; Nguyen & Santagata, 2020; Scherr & Robertson, 2015). Such activities have given rise to new symbols, signs, codes, and even languages normally associated with the computer science (CS) discipline and used by students to understand scientific phenomena (Abrahamson & Sànchez-Garcià, 2016; Aksit & Wiebe, 2020; Chandrasekharan & Narsessian, 2015).

Not only can computational modeling activities improve the learning of scientific concepts, Berland and Wilensky (2015) and Aksit and Wiebe (2020) showed that such activities also enhanced another essential science and engineering practice, called computational thinking (CT) skills. Weintrop et al. (2016) argued that systems thinking is part of CT and that, hypothetically, understanding of both constructs would increase after participating in a computational modeling activity. The purpose of this study is to examine the impact of learning about food webs in a computationally rich environment on middle school students' conceptual understandings of food webs, systems thinking skills, and CT skills. This strategy is a strong example of promoting 3D learning, an instructional design that combines the three dimensions of

disciplinary knowledge, practice, and crosscutting concepts (NGSS Lead States, 2013). In addition, the current study directly explores Weintrop et al.'s (2016) theoretical prediction of a hierarchical relationship between systems and computational thinking skills.

Conceptual Framework and Literature Review

Systems Thinking

Thinking about systems or systems thinking is a crucial habit of mind for many science fields that centers on the understanding of complex systems and problems and the properties and behaviors of the systems (Mambrey et al., 2020a). In the General System Theory, von Bertalanffy (1973) derived the idea of systems thinking from many different fields, including systems in natural and social sciences, by seeing the commonality in methodological thinking. Given this broad definition of systems thinking, research in science education has identified many different types of thinking processes that fall under the umbrella of systems thinking, such as thinking in levels (Wilensky & Resnick, 1999; Boersma et al., 2011), causal reasoning (Hmelo-Silver et al., 2000; Jacobson & Wilensky, 2006), mechanistic reasoning (Hmelo-Silver et al., 2017; Krist et al., 2019; Russ et al., 2008), structure-function-behavior (Hmelo-Silver et al., 2009; Lavi et al., 2020), dynamic thinking (Maani & Mharaj, 2004; Hrin et al., 2017), cyclic thinking (Ben-Zvi Assaraf & Orion, 2005; Hrin et al., 2017), and interdisciplinary thinking (Stratford et al., 1998; You et al., 2017). Mambrey et al. (2020a) offered a more operationalized definition: "as a conceptual skill in which superordinate principles of complex systems are taken into account when understanding and predicting the interplay and function of their elements" (p. 3).

Although researchers have been using different terms for systems thinking, there is a significant commonality that makes those terms or thinking processes part of systems thinking,

which is the involvement of complex problems or systems, as reflected in Mambrey et al.'s definition. Mehren et al. (2018) characterized complex systems in which: (1) systems that model the complex realities, (2) showing the complexity of structure, function, and behaviors, including the revelations of linear and nonlinear interactions among the components and emergent effects, (3) the systems interact with other systems, and (4) the patterns of how the systems work are not explicitly targeting towards a single goal. Sweeney and Sterman (2007) conducted a qualitative study on teachers' and students' systems thinking skills using natural and social systems scenarios (e.g., predator and prey, practice and performance, etc.). They found that students often employed more simple causal thinking, which made them unable to attend to more complex cause-and-effect thinking. This points to the need for more understanding of how to aid students in developing higher order systems thinking skills.

Many studies have been conducted to identify the components of systems thinking so that researchers can develop interventions aiming specifically at those components. Through a series of studies on elementary, high school, and university students learning earth system and environmental system, Ben-Zvi Assaraf and Orion's groups (Ben-Zvi Assaraf & Orion, 2005a, 2010a, 2010b; Batzri et al., 2015) developed a systems thinking framework called System Thinking Hierarchic or STH. According to the STH framework, systems thinking composed of eight components of skills, which are: "(1) the ability to identify the components of system and processes within the system, (2) the ability to identify relationships among the system's components, (3) the ability to organize the systems' components and processes within a framework of relationships, (4) the ability to make generalizations, (5) the ability to identify dynamic relationships within the system, (6) understanding the hidden dimensions of the

system, (7) the ability to understand the cyclic nature of systems, (8) thinking temporally [including] retrospection and prediction" (Ben-Zvi Assaraf & Orion, 2010a, p. 1255).

Researchers have used the STH framework to examine students' systems thinking when learning scientific concepts (Bergan-Roller et al., 2018; Lee et al., 2019; Scherer et al., 2017). Lee et al. (2019), in a study on pre-service science teachers' use of systems thinking in the water cycle, found that it is not enough to define leveling in systems thinking based on the cognitive skills mentioned above. In response to the structural issue of systems thinking, Hokayem and colleagues (Hokayem, 2016; Hokayem et al., 2015; Jin et al., 2019) conducted a series of studies on elementary students' systems thinking skills within the ecosystem concepts. They found that systems thinking should be conceptualized within learning progression theory, in which students' systems thinking skills gradually increase. Hokayem concluded that the components of systems thinking are not hierarchically organized, rather side-by-side. To synthesize the findings suggested by Hokayem et al. and Lee et al., Mehren et al. (2018) and Mambrey et al. (2020a) conceptualized that systems thinking consists of three components – system organization, system behavior, and system-adequate intention to act or modeling–, and these components are organized side-by-side, but they are progressed in stages (e.g., Stage 1, Stage 2, Stage 3). The authors operationalized this conceptualization by developing a theoretical foundation and assessments measuring students' systems thinking skills in geography and food webs.

One aspect that Mehren et al. (2018) and Mambrey et al. (2020a) did not examine in their study was using their assessments to examine the impact of an intervention of students' systems thinking. Such an intervention study would be an important step in both validating the framework and contributing to the ongoing validation of the assessments. Mambrey et al. (2020a) suggested using model-based science instruction to operationalize this framework. The

use of model-based science instruction is one effective instructional strategy to improve students' systems thinking skills (Berland & Wilensky, 2015; Bodzin, 2011; Hmelo-Silver et al., 2017; Wilkerson-Jerde & Wilensky, 2015). For example, Berland and Wilensky (2015) examined the impact of two-week-long robotics activities on students' complex systems and computational thinking skills. They compared physical and virtual robotics activity. They found that there was a significant increase in complex systems thinking in the virtual robotics condition. A logical next step would be to investigate whether Mehren et al. and Mambrey et al.'s systems thinking framework and assessment works well in model-based science instruction, especially when students are involved in scientific computational modeling activities.

Learning Food Webs Concepts

Research has shown that students struggle with learning ecological concepts at the K-12 level. Hogan and Fisherkiller (1996) studied fifth and sixth graders' reasoning about food webs, decomposition, and the carbon cycle. They found that students could not recognize that carbon is recycled in the ecosystem and decomposition is the key component in food webs and nutrient cycling. Similarly, other researchers found the same results regarding students' understanding of food webs (e.g., Eberbach et al., 2012; Eilam, 2012; Hogan, 2000; Reiner & Eilam, 2001), particularly the misconceptions they have around the predator/prey relationship, as described in the introduction above. Through studying secondary level students, these authors found that at this grade level students were able to recognize the phenomena of predator and prey eating order in a food web, but could not see other key relationships, such as energy transmission. Reiner and Eilam (2001), who studied ninth-graders, called this reasoning phenomenon incoherent but consistent, meaning that students incoherently relate the process of feeding to other components of food webs, as well as other matters such as energy. Yet, they are consistent with applying the

concept of eating order. Furthermore, Eberbach et al. (2012) offered an alternative explanation to this reasoning based on their qualitative study on 12 middle school students' understanding of the ecosystem. They analyzed it using a systems-thinking approach—structure-behavior-function framework. They found that understanding the multiple relationships in many levels within a food web and ecosystem is crucial for a more coherent conceptual understanding of food webs.

Anderson et al. (2009) and some other researchers (Hokayem & Gotwals, 2016; Jin et al., 2019; Mohan et al., 2009) took a learning progression approach to study students learning of food web or ecosystem concepts, intending to provide a fine-grain understanding of what students could learn in each level of education. In general, studies on learning progressions of the ecosystem concepts in fourth to twelfth grades identified four levels of progression (Jin et al., 2019; Mohan et al., 2009), starting from level one and two, where students only account for the agents within the system, as opposed to the processes to level four, that puts more emphasis on processes. For middle school students, according to Mohan et al. (2009) it is expected that they would exhibit either level one or two. In addition to the characteristics of levels one and two mentioned earlier, students in these levels tend to reason by utilizing the concepts of natural vitality and tendency, or more teleological (Leach et al., 1996), such that rabbits eat grasses to become stronger or to grow.

Although findings from learning progression areas provide a clear distinction on how students learn food webs and progress throughout their educational years, many researchers studied how appropriate instructional strategies could aid students in seeing more sophisticated relationships in an ecosystem (e.g., Grotzer & Bell, 2003; Lehrer & Schauble, 2012; Nguyen & Santagata, 2020). Notably, using scientific modeling and computer simulations approach, Lehrer and Schauble (2012) saw development in K-6 students' conceptual understanding of the

interdependency in an ecosystem, such that the students could recognize more interactions of abiotic and biotic factors. Consistent with this finding, Dickes et al. (2016) qualitatively studied elementary students' learning of ecosystem concepts through agent-based modeling; analyzing the data from a mechanistic reasoning perspective. The authors found out that their participants developed a more advanced conceptual understanding of interdependence in an ecosystem after a two-week-long intervention. More recently, Nguyen and Santagata (2020) conducted an experimental study of sixth graders learning about decomposition in a food web. The authors found that students had an increase in their systems thinking skills when they learned in both a computational modeling condition and a paper and pencil-based modeling, however the former had a larger effect size than the latter. These studies have suggested the potential of computer-based modeling in aiding students in learning food web concepts and systems thinking. Yet, none of the previous studies emphasize the impact on the emerging and essential science and engineering practice, computational thinking, that many have found is developed through engaging in computational modeling activities (e.g., Aksit & Wiebe, 2020; Yin et al., 2020). Through a qualitative study analyzing science education-related documents and observing classroom activities, Weintrop et al. (2016) argued that, hypothetically, systems thinking is part of computational thinking. However, this relationship has been understudied in the literature, especially when both constructs are measured repeatedly, such as before and after an intervention.

Computational Thinking (CT)

Literature has identified two different yet related conceptualizations of CT – one that is highly related to computing and computer science contexts (e.g., Brennan & Resnick, 2012, 2018; Denner et al., 2012; Guzdial, 2008; Lye & Koh, 2014; Weintrop et al., 2016), while the

second one focuses more on CT as a set of general problem-solving skills (e.g., Selby & Wollard, 2013; Shute et al., 2017; Yadav et al., 2014). Wing defined CT as "an approach to solving problems, designing systems, and understanding human behavior that draws on concepts fundamental to computing" (Wing, 2008, p. 3717). Wing then further clarified her definition of CT as "the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent" (Cuny et al., 2011, p. 20). Aho (2012) simplified Wing's definition by focusing more on the formulation of problems in which the "solutions can be represented as computational steps and algorithms" (p. 832). In line with Aho's definition, Holbert and Wilensky (2011) defined CT as "the ability to translate or encode phenomena (real or imagined) into representations that leverage computational power" (p. 110). Lee et al. (2014) also offered a quite similar definition in which they defined CT as "a way that humans think and that the ability to harness the power of computers" (p. 65). Additionally, in the United Kingdom (UK), as part of introducing computing nationally, the Royal Society (2012) defined that CT "is the process of recogni[z]ing aspects of computation in the world that surrounds us, and applying tools and techniques from Computer Science to understand and reason about both natural and artificial systems and processes" (p. 29). These definitions were then adopted and used in the K-12 Computer Science Framework (2016).

There continues an ongoing discussion around a unified CT definition, as seen in the many published systematic review studies in the past few years (e.g., Lye & Koh, 2015; Shute et al., 2017; Tang et al., 2020). One of the notable reviews is Shute et al.'s (2017), where they reviewed seventy published studies on CT in K-16 settings. Their review produced a definition of CT that includes the applicability of CT in different contexts and academic areas and does not

depend on the solution being machine/computer readable. They defined CT as “the conceptual foundation required to solve problems effectively and efficiently (i.e., algorithmically, with or without the assistance of computers) with solutions that are reusable in different contexts” (p. 151). For this study, the different contexts that Shute et al. refer to are conceptualized as scientific concepts and the phenomena that students encounter both in classroom and daily-life settings.

Research on Computational Thinking in Science Education

The inclusion of CT into science curricula has been forwarded by the use of model-based science learning, particularly with the implementation of constructionism in learning that was popularized by Papert (1980). According to constructionism, learning is designing and building objects that embody a content domain’s fundamental ideas. DiSessa (1979) argues that these objects can be in the form of “microworlds” that represent “the powerful logical structure of the subject, but which at the same time mesh properly with the cognitive reality of human beings” (p. 239). Papert (1980) asserted that students involved in constructionist learning activities typically use an authoring environment where they design and program artifacts that constitute the underlying structures and rules of a phenomenon. Teachers and students who are involved in CT-integrated science learning (also called *computationally rich science activities*) typically use the representation of scientific concepts in the form of computational models rather than physical models (e.g., Aksit & Wiebe, 2020; Dickes et al., 2016), though some studies use the combination of both (e.g., Wade-James et al., 2018). Research has also indicated that CT-integrated science instructions, either when students only use computational models or build them from scratch, could improve students’ understanding of scientific concepts (e.g., Aksit & Wiebe, 2020; Peel et al., 2019) and advance their scientific reasoning (e.g., Dickes et al., 2016;

Irgens et al., 2020; Pallant & Lee, 2015; Pierson & Clark, 2018), representational fluency (Moore et al., 2020), and epistemic agency (Irgens et al., 2020; Pierson et al., 2020). Moreover, some also found that the instructions influenced students' social aspects of learning (e.g., Pierson & Clark, 2018; Pierson et al., 2020) and science interest and motivation (e.g., Leonard et al., 2016; Wagh et al., 2017).

Integrated CT-Science learning has often taken the form of agent-based modeling, a specific form of computational modeling. Students design each component or actor in the model, especially their behaviors, from scratch by utilizing a computer programming environment (Sengupta et al., 2013). Wagh et al. (2017) argued that agent-based modeling activities through computer programming could bridge the idea of constructionism and inquiry-based science. Wagh et al. (2017) found that tenth grade students' tinkering with computer code provided disciplinary engagement in inquiry-based science in two ways: computational and conceptual engagement. Computational engagement refers to using codes as a representational medium to answer questions. In contrast, conceptual engagement is defined as noticing and explaining components that can result in changes in the modeled phenomena. Therefore, learning science through modeling scientific phenomena in a computational programming environment has been found to improve both CT and conceptual understanding of science. For instance, Aksit and Wiebe (2020) investigated the impact of a week-long computational thinking and modeling activity in a regular seventh-grade science classroom on students' learning of force and motion and computational thinking skills. They used a one-group experimental study without a control group. They found that participating in the activity increased students' conceptual understanding of force and motion and computational thinking skills with a large effect size.

Interestingly, CT practices in science learning without any computers or computational models (called unplugged activities) have also been found to improve students' science learning. Peel et al. (2019), examined the impact of using CT concepts and practices on high school honor biology students' conceptual understanding of natural selection. They found significant scientific concept use gains along with a significant decrease in misconceptions. Regarding the relationship between CT and natural selection factors, they found a strong correlation between the uses of CT-related *branching* or conditional logic, i.e. "[c]hoosing a path, if/then/else statements…example: Driving through traffic lights: IF the light is green, THEN you go, ELSE you slow down and stop" (Peel & Friedrichsen, 2018, p. 22) to explain differential survival. Last, the authors also found that students expressed that the generalized algorithmic explanation helped them make sense of natural selection.

Current Study's Purposes and Research Questions

The above literature review has provided insights into how students, particularly middle school students, learn food web concepts, especially from systems thinking perspectives. Literature has also indicated that both computer-based and paper-based modeling can significantly impact students' conceptual understanding of food web and systems thinking skills. What has been less studied is how these two instructional approaches influence students' CT skills. Some researchers such as Weintrop et al. (2016), hypothesize that systems thinking is hierarchically part of CT, while NGSS (NGSS Lead States, 2013) categorize them in two separate dimensions: crosscutting concepts, and science and engineering practices. The current study compares the impacts of these two modeling activities on students' computational and systems thinking skills as well as conceptual understanding of food webs through conducting a quasi-experimental study. In addition, the purpose of this study is to gain more insight into the

relationship between CT and systems thinking. The following research questions guide the current study:

1. What are the impacts of computational modeling activities on students' systems thinking when learning about food web concepts?
2. To what extent do scientific computational modeling activities improve students' CT skills?
3. To what extent does systems thinking relate to computational thinking?

Methods

Research Design, Context, and Participants

The current study adopted a two-condition quasi-experimental design. A total of 751 parent-consented and assented seventh-grade Indonesian students participated in this study. These students were from nine different schools taught by 11 teachers. Due to the COVID-19 global pandemic, all students were engaged in a series of online learning, either synchronous or asynchronous and thus, the attrition rate was quite high. For this study's purpose, only students who attended class and completed worksheets for three out of four activities were included in the current study. This resulted in a total of 365 students in the final dataset. Results from independent sample *t*-test on these students' end-semester science and math scores indicated that students in final dataset had significantly higher science and math scores than those who were excluded, but with a small effect size (t = -2.47, p = .014, d = 0.23; t = -2.17, p = .031, d = 0.21, respectively). Also, regarding socioeconomic level, the final dataset contained a significantly lower number of students at a low SES level and a higher number of students in high SES level than those in the dropped dataset (X^2 = 34.10, p < .001). However, there was no difference based gender proportion in the two datasets (X^2 = 0.02, p = .891). This was expected as many studies

have also found that students who tended to complete and actively participate in an online learning environment tend to be high achieving students because they tended to have higher persistence and locus of control than the dropped students (Deschacht & Goeman, 2015; Lee, Choi, & Kim, 2013). However, we acknowledge this as one of the limitations of the current study.

Regarding demographics, 60% of students were from high SES level, 26% medium SES level, and 14% low SES level. Please see the Methods subsection on Covariates for a description of how SES was calculated. There was an equal number of male (49%) and female (51%) students. These students were then assigned purposefully based on whether they owned a computer machine or not to either synchronous computational modeling condition or asynchronous paper-based pictorial modeling condition. Due to this design, in the analysis students' SES levels were controlled.

Curricular Context

Both the control and experimental conditions covered the same food-web concepts, with computational modeling used in the experimental condition and the control condition using paper-based pictorial modeling. All activities were embedded as part of the regular teaching and learning activity for an ecosystem unit in science classrooms, lasting for 4 x 60 minutes. The following sections provide a detailed description of each condition.

Experimental Condition: Computational Modeling Activity

The computational modeling activity for learning food web material consisted of four different activities, built based on the Use-Modify-Create (UMC) framework (Lee et al., 2011). The researcher trained all the teachers who agreed to participate in this study. The Cellular agent-based programming environment (Meyer et al., 2012), an extension of block-based

programming, Snap! (Garcia et al., 2015), was used as a platform on which students built their computational models (Houchins et al., 2021; Lytle et al., 2019). In the activity, students learned the concept of energy transfer within a simple food web model. The computing-infused activity enabled students to explore energy transfer in a food web and thus understand the food web concepts beyond simple eating order. Even though each activity was designed to build upon the prior one, students started with the same code at the beginning of each activity to reduce absence and incompletion effects.

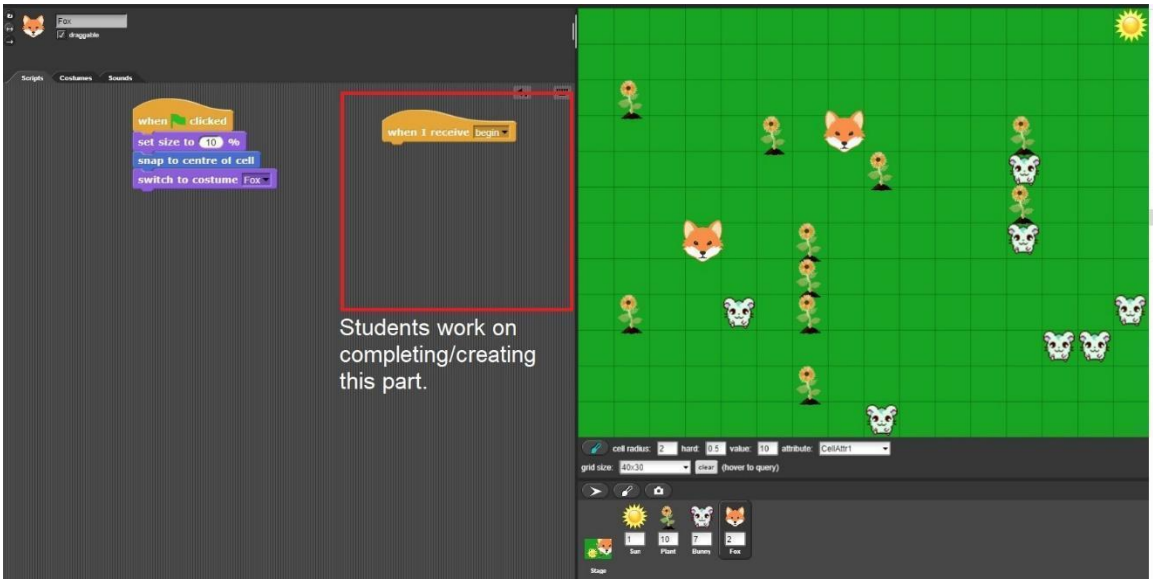
Activity 1: Unplugged. The first and only activity that did not involve a computer is referred to as "unplugged." Students learned the basic concepts of food webs in this activity, including their definition and components, primary and secondary consumers, and the energy transfer within the food web system. At the end of the activity, the teachers and the students completed a worksheet on which they described the agents' behaviors that they would have in the computational model. Their descriptions were in the form of pseudocodes. This served as a transitional representation of the core ideas and a stepping stone to creating block-based codes.

Activity 2: Use. The focus of the second activity was understanding the relationship between the "Plant" agents (the producer) and the "Sun" agent that provides energy to the plants so they can grow. Students did not build their codes; instead, they interacted with a completed computational model, inspected the output simulation, and read through the working code. This *Use* activity was intended to familiarize students with the block-based programming environment and the weather conditions that influence plants' growth. The teachers led students in exploring the working programming codes by changing the default input or initial intensity of solar energy, the cutoff conditions (i.e., energy needed for transition or growth), and the amount of energy lost through growing or transitioning. Students then recorded on a worksheet all the

changes they had in their model. This worksheet made students focus on how the changes influence the plants' growth speed, from seed to flower state.

Figure 1

A Food Web on Cellular Platform Showing Plants, Bunnies, and Foxes on the Grid Background for Activity 4/Create (Note: Students Used the Bahasa Version)



Activity 3: Modify. In the third activity, students modified the "Bunny" agent, the primary consumer model. Students received the "Bunny" code at the beginning of the activity. However, the bunny's given behavior did not align with the scientific model—e.g., bunnies never eat the flower when they are hungry, flowers change to a wrong state after a bunny eats them. Students modified existing programming codes to make them align with the scientific model that was planned and discussed in the unplugged activity to complete this Modify activity.

Activity 4: Create. The last activity focused on modeling the "Fox" agent, the secondary consumer. The students were asked to create a script containing codes representing the Fox. They modified the code for Bunny to appropriately interact with this newly introduced agent.

Once complete, as Figure 1 illustrates, students received a task to change some functionality in their programming codes and compare how their new simulation model acted differently from the previous one.

Control Condition: Paper-based Pictorial/Drawing Modeling

Students in the control group learned the same food web concepts as those in the experiment group. Students in the control group took part in four online module activities where they worked on worksheets and drew the food webs. They opened and went through the modules and submitted their drawings and diagrams using their smartphone at a scheduled time for science class. Students were given 60 minutes to complete the activity/module.

Activity/Module 1. Like the experimental group Activity 1, students learned the basic concepts of food webs, including their definition and components, primary and secondary consumers, and the energy transfer within the food-web system. They created a representation of the scientific model as a series of logical conditionals (e.g., If...then). Students were also asked to draw a simple food web that shows interrelationships between biotic and abiotic factors in a food web. Figure 2 presents some of the students' drawings.

Activity/Module 2. In Activity 2, students drew the connection between the sun's different conditions and transmitted solar energy that plants receive. They were engaged in an activity to predict the relationship between solar energy and plants' growth, based on their understanding of the two components and their drawings. Figure 2 shows some examples of students' works in this module.

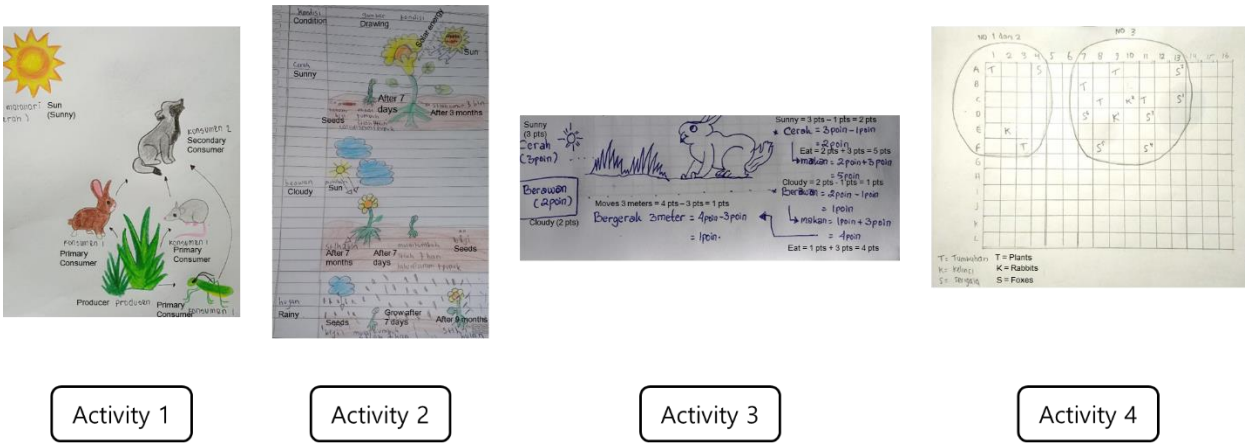
Activity/Module 3. Students then were involved in an activity where they added a primary consumer to their model and received a set of questions and scenarios related to the primary consumer's behavior. The students were expected to find the relationship between the

different conditions of sun, producer, and primary-consumer behavior. Figure 2 depicts students' artifacts from this activity.

Activity/Module 4. In the final activity, students worked on a worksheet that contains 7 x 5 grid boxes. They were asked to add a secondary consumer to their drawing. Students took part in an activity where they used the grid to answer the questions about the scenarios they received. These questions about the scenarios consisted of all components they had drawn and asked them to explore their relationships.

Figure 2

Students' Artifacts for Each Activity in Paper-Based Pictorial/Drawing Modeling Condition



Data Sources

Computational Thinking Assessment.

Students' CT skills were measured by using Wiebe et al.'s (2019) CT assessment. This assessment consisted of 25 multiple-choice questions that comprise six puzzle-like items from Bebras challenges (Bebras, 2017) and 19 items from Roman-Gonzalez et al.'s (2017) CTt that uses a generic block-based computer programming language as the context of the questions. The assessment is appropriate for this study, as it was validated for middle-grade levels and can be completed in a short class period. The assessment was initially developed in English and

translated to Indonesian through a back-translation process (Brislin, 1970; Cha et al., 2007). Students in both conditions took the assessment for about 30 minutes before and after the intervention. Before running the primary analysis, we ran Rasch modeling to re-validate the instrument and found two misfitting items (MNSQ > 1.30). We then removed these two items in further analysis. The expected-a-posteriori reliability of the final version's plausible values and separation reliability were .654 and .984, respectively. We acknowledged this relatively moderate EAP/PV reliability value; however, we argued that we could still use the instrument in this study, given the items were psychometrically sound to measure students' CT skills (Supporting Materials A). The low-reliability value is possibly due to the broad range of students' CT skills, and thus the items in this instrument could not locate all the students' CT skills (Bond et al., 2021).

Systems-Thinking-Embedded Food Web Assessment

This study utilized Mambrey et al.'s (2020a) 36 item systems-thinking-embedded food web assessment to measure students' systems thinking skills and conceptual understanding of food webs. This assessment was developed based on Mehren et al.'s (2018) conceptualization of systems thinking. The assessment measured three components of systems-thinking skills, namely, systems organization, system behavior, and system modeling. Tasks related to system organization include the ability to identify pertinent populations in the ecosystem and "the existing (predator-prey) relations" (Mambrey et al., 2020a, p. 8). For system behavior, students work on tasks that require them to describe the results of changes in population sizes. In the last component, system modeling, students are asked to work on problems related to the development of "prognoses and system regulation" (Mambrey et al., 2020a, p. 8). As systems thinking is related to understanding the complexity of systems, Mambrey et al. developed the instruments by

combining these three component skills with sophistication levels in food-web systems, resulting in nine combinations of item types. These levels of sophistication include

1. Direct relation: In this level, students exhibit reasoning about simple systems of food webs based on only monocausal relations, such as predator and prey;
2. Indirect linear relations: In this level, students show reasoning about moderately cross-linked systems of food webs, based upon linear relations;
3. Indirect complex relations: In this level, students show reasoning in highly cross-linked systems of food webs.

Similar to the approach we took on the CT assessment, a back-translation process was done for this assessment (Brislin, 1970; Cha et al., 2007). Furthermore, multidimensional Rasch modeling was performed to validate the assessment with our data. The results suggested removing three items from the systems organization dimension and one item from the systems behavior dimension due to item misfitting issue (MNSQ values > 1.30). We then proceeded with 32 items for the subsequent analysis. This assessment's separation reliability was .996, and the EAP/PV reliability values for each dimension were .847, .848, and .845, respectively, for systems organization, behavior, and modeling. The complete validation results are available in Supporting Materials B.

Covariates

Students' SES information was controlled, given that the condition assigned to each student was based on the characteristics of the students. Students' socioeconomic (SES) levels were collected by asking them to rate their family's income, computer/laptop possession, and parents' educational levels. Polytomous Latent Class Analysis (poLCA; Linzer & Lewis, 2011) in R was run to cluster students' SES levels based on these constructs. The results showed that

the 3-clusters solution was the best fit for our data, indicating low, medium, and high SES levels (for detail, see Supporting Materials C). Students were also asked to voluntarily provide their latest score science and mathematics scores reported in their end-semester grade reports. Together, these variables were controlled to examine the impact of the intervention on students' systems thinking and CT skills.

Daily Exit Ticket

After students completed each activity, they took a short formative assessment called "Daily Exit Ticket." In this formative assessment, we asked what students experienced and learned from each activity through this open-ended question: "What did you learn today?." We explicitly asked students to provide at least two complete sentences in their answers to gather richer data.

Data Analysis

Quantitative Data

Given that the collected data were both repeated measures and nested data, which are likely to be unbalanced, Raudenbush and Byrk (2002) suggest using multilevel modeling (MLM) to analyze such data. MLM is able to present the intra-individual variability, which is the impact of the intervention on students' variability (Raudenbush & Bryk, 2002). A fully unconditional model was run to test whether there was enough within- and between-student variance to continue with the conditional MLM models (Raudenbush & Bryk, 2002). The unconditional model consisted of no predictors and generated within-person (σ^2) and between-person (τ_{00}) variability. The estimates can be used to calculate the intraclass correlation coefficient (ICC) [$\rho = \tau_{00} / (\tau_{00} + \sigma^2)$], which shows the amount of between-person variance in systems thinking and CT

(dependent variables). Furthermore, an approach suggested by Snijders and Bosker (2011) was used to calculate within- and between-student variability in the conditional models.

MLM was performed using the PROC MIXED package in SAS (1997). This study investigated the impact of the interventions on students' systems-thinking and CT skills (Level 1). We were also interested in determining whether the type of condition impacts the changes, if any, in students' CT and systems thinking scores when controlling for students' SES level (Level 2). At Level 2, we added students' mathematics and science scores to examine whether there was a difference in changes when students had an average score on these subjects. All of these were tested using the following formulas:

Level 1: Time	$\text{Systems Thinking}_{it} = \beta_{0it} + \beta_{1it}\,(\text{Test Occasion}) + r_{it}$	(1)
Level 2: Student	$\beta_{0i} = \gamma_{00} + \gamma_{01}\,(\text{Condition}) + \gamma_{02}\,(\text{SES}) + \gamma_{03}\,(\text{ScienceScore}) + \gamma_{04}\,(\text{MathScore}) + u_{0i}$	(2)
	$\beta_{1i} = \gamma_{10} + \gamma_{11}\,(\text{Condition}) + \gamma_{12}\,(\text{SES}) + \gamma_{23}\,(\text{ScienceScore}) + \gamma_{14}\,(\text{MathScore})$	(3)

Equation (1) specifies within-student relationships between the repeated measures (pre-posttests). The intercept β_{0it} is the expected score of systems thinking for student *i* on the pretest and when the science and mathematics scores are on the average (grand mean centering was applied to science and mathematics scores). The first slope, β_{1it} is the anticipated change in the systems thinking score (or other dependent variables such as CT, systems organization, behavior, or modeling) obtained before and after the intervention (test occasion), which is the change associated with moving from the pretest (pretest coded as 0) to the posttest (posttest coded as 1). The error term r_{it} represents the amount of deviation around the mean of systems thinking score. We named this as Model 1 without any covariates. The intercept and slopes in Level 1 become the outcomes at the student level or Level 2. Equations 2 and 3 include the

condition/intervention, students' SES level, and science and mathematics scores. The intercept γ_{00} denotes the average of students' systems thinking scores when every predictor is at 0 (average science and mathematics scores). Equation 3 yields γ_{10}, the average of students' systems thinking score (or other dependent variables) before and after the intervention, and γ_{11}, which tests whether there is an effect of the intervention on the relationship between students' systems thinking score and test occasion, controlling for the students' SES. γ_{11} was used to answer the first and second RQs. We used the same equations for the three systems thinking dimensions and CT.

Repeated-measures correlation tests were run using the rmcorr package in R. These tests were intended to answer the third RQ regarding the relationship between CT and systems thinking. Repeated-measures correlation is "a statistical technique for determining the common within-individual association for paired measures assessed on two or more occasions for multiple individuals" (Bakdash & Marusich, 2017).

Qualitative Data

Students' responses to the open-ended question in daily exit tickets were analyzed qualitatively. The coding scheme developed by Mambrey et al. (2020b) was utilized to analyze the data. Mambrey et al. (2020b) developed this coding scheme by following a theory-based coding-scheme procedure (Schreirer, 2012). They developed this coding scheme based upon systems-thinking theory and factors that influence or function in the students' learning process for food webs. Two additional codes related to computer science/CT were added after a round of open coding (Strauss & Corbin, 1998). A trained second coder independently coded 20% of randomly selected data to generate inter-coder reliability. The final coding scheme had a satisfactory inter-coder reliability value k = .882 (Cohen, 1960). When first and second coders

disagreed, they met and discussed their disagreements until they reached a consensus, which might lead to the modification of the codebook. The final code is presented in Table 1.

Students' responses were analyzed using Epistemic Network Analysis (ENA, Shaffer & Ruis, 2017). ENA is a quantitative ethnographic approach to model the structure of co-occurrences and connections of qualitative data codes. Even though ENA has been traditionally used to model the structure of conversation, we followed the approach used by Siebert-Evenstone and Shaffer (2019) that modeled the structure of connections between crosscutting concepts and practices across disciplines in NGSS using performance expectations (PEs) as their data. They defined conversation as lines of PE. Similarly, in the current study, we defined conversation as students' responses on each activity day. ENA generates a network visualization of codes (corresponding to nodes) listed in Table 1 and connections between codes (corresponding to edges). The larger the nodes indicate, the more frequent codes mentioned by students, and the thicker and more intense color of the edges, the more co-occurrences between two codes. The network visualization shows the differences and similarities of students' perceptions in the two conditions. It is important to note the relative nature of the visualization rendering. Thus, the line and node size indicate occurrence rates relative to the other condition, not an absolute value. These ENA visualizations are beneficial for triangulation purposes to substantiate evidence from quantitative results in answering the three research questions.

Table 1

A Coding Scheme for Students' Responses in Daily Exit Ticket Data

Category	**Description**	**Example**
System Organization (*k* = .804)	Students identified the component of food webs (monocausal, linear, or complex)	"...rabbits eat plants" [SMP2B7G04] "rabbits eat plant and then rabbits are eaten by fox" [SMP2B7E07]

Table 1

A Coding Scheme for Students' Responses in Daily Exit Ticket Data (Continued…)

Category	Description	Example
System Behavior (*k* = .881)	Students expressed the interrelation and interaction among components of food web (monocausal, linear, or complex)	"...solar energy has to do with the relationship between producer, consumer 1, and consumer 2" [SMP2B7C10] "...I learned how producer and primary consumer interacted in different weather conditions" [SMPIS7C04]
System Modeling (*k* = .842)	Students mentioned about the predictions of if something happens within the food web and how it affects the food web in the future (monocausal, linear, or complex)	"...to predict the effect of solar energy on organisms...and the relationship between solar energy, producer, consumer 1, and consumer 2" [SMPLU7B04]
Biological Concepts Food Webs (*k* = .944)	Students indicated that they learned food web concepts	"I learned about food webs and food chain" [SMP2B7G13] "I learned about producer, primary consumer, and secondary consumer" [MTSAS7D07]
Biological Concepts Photosynthesis (*k* = 1.000)	Students mentioned that they learned the concept of photosynthesis	"I learned about coding, plant growth, and photosynthesis" [SMPDT7C21]
Biological Concepts Energy (*k* = .862)	Students expressed that they learned the concepts of energy	"I learned the process of energy transfer within a food web" [SMPLU7A21]
Representational Features (*k* = .834)	Students identified the component of modeling, such as when they mentioned words model, SNAP, animation, and simulation.	"Today, I learned food web using simulation through SNAP" [SMPDT7C27]
Pseudocode (*k* = .933)	Students indicated that they learned the concept of pseudocode	"...I learned script and pseudocode" [SMP2B7A17]

Table 1

A Coding Scheme for Students' Responses in Daily Exit Ticket Data (Continued…)

Category	Description	Example
Coding and Programming (*k* = .836)	Students mentioned that they learned how to code/program, including when they expressed the word "block" as it tended to refer to code block.	"I learned programming language and coding about food web" [SMPLU7C40]

Findings

RQ 1. Impacts on Systems Thinking

We organized the findings for systems thinking by reporting each of the systems thinking components' results, followed by the overall systems thinking skills score. For each of the components, we first examined the results from a fully unconditional or null model to determine the within- and between-students variability in each of the systems thinking components that were also used to determine whether MLM could be used for further analysis. Then, we reported the three conditional models outlined in the Methods section to investigate the impacts of computational and paper-based pictorial modeling on students' systems thinking.

System Organization

The results from a fully unconditional model showed that 66% of the variability in system organization scores was between-student ($\tau_{00} = 1.95$, $z = 10.03$, $p < .001$), and 34% was within-student ($\sigma^2 = 0.99$, $z = 12.08$, $p < .001$). Thus, these results indicated that there was sufficient variability to proceed with MLM. The conditional models were then run and the results presented in Table 2.

Based on Table 2, multilevel modeling results showed a significant increase ($p = .001$) in students' systems organization scores from pre- to post-test in both computational and paper-

based modeling conditions (Model 1). SES and academic scores were then incrementally added to Model 2 and Model 3. We found a significant relationship between students' mathematics scores on system organization scores (Model 3, $p < .001$). However, we found that the substantial increase from pre- to post-test in students' system organization scores became insignificant (Model 3, $p = .447$). Model 3 of system organization accounted for 25% and 20%, respectively, of the within- and between-student variability in systems organization scores. Figure 3 visualizes the results after controlling for students' SES and when they have average academic scores.

Table 2

Unstandardized Coefficients (and Standard Errors) of Multilevel Models of Systems Organization

Effect	Parameter	Model 1: Condition Effect	Model 2: Controlled SES	Model 3: Controlled SES + Academic Scores
Systems Organization, β_0				
Intercept	γ_{00}	-0.59***(0.15)	-1.57***(0.33)	-1.00** (0.33)
Condition (Experimental)	γ_{01}	0.85***(0.19)	0.39 (0.23)	0.41* (0.19)
SES	γ_{02}		0.51** (0.16)	0.18 (0.13)
Science Score	γ_{03}			-0.00 (0.01)
Mathematics Score	γ_{04}			0.03***(0.01)
Intervention slope, β_1				
Time (Post)	γ_{10}	0.42** (0.13)	0.34 (0.29)	0.44 (0.31)
Condition (Experimental)	γ_{11}	-0.02 (0.16)	-0.05 (0.21)	-0.11 (0.21)
SES	γ_{12}		0.04 (0.14)	-0.01 (0.14)
Science Score	γ_{13}			0.01 (0.02)
Mathematics Score	γ_{14}			0.01 (0.01)
Random Effects				
Variance components				
Between-student (τ_{00})		1.834***(0.18)	1.738***(0.18)	1.465***(0.16)
Within-person fluctuation (σ^2)		0.912***(0.08)	0.913***(0.08)	0.896***(0.08)

Notes: no asterisk $p > .05$, *$p < .05$, **$p < .01$, ***$p < .001$; Reference group = Pretest, Control Condition, and average Science and Mathematics Scores

System Behavior

The results from a fully unconditional model showed that 69% of the variability in system behavior scores was between-student ($\tau_{00} = 1.91$, $z = 10.31$, $p < .001$), and 31% was within-student ($\sigma^2 = 0.87$, $z = 12.07$, $p < .001$). These results also indicated sufficient variability in students' system behavior scores to proceed with MLM. The conditional models were then run, and the results are presented in Table 3.

Table 3

Unstandardized Coefficients (and Standard Errors) of Multilevel Models of Systems Behavior

Effect	Parame-ter	Model 1: Condition Effect		Model 2: Controlled SES		Model 3: Controlled SES + Academic Scores	
Systems Behavior, β_0							
Intercept	γ_{00}	-0.13	(0.14)	-0.97**	(0.32)	-0.51	(0.33)
Condition (Experimental)	γ_{01}	0.96***	(0.18)	0.58*	(0.23)	0.58**	(0.23)
SES	γ_{02}			0.43**	(0.15)	0.25	(0.15)
Science Score	γ_{03}					-0.01	(0.01)
Mathematics Score	γ_{04}					0.04***	(0.01)
Intervention slope, β_1							
Time (Post)	γ_{10}	0.31*	(0.12)	0.37	(0.29)	0.57	(0.30)
Condition (Experimental)	γ_{11}	-0.25	(0.16)	-0.21	(0.20)	-0.33	(0.20)
SES	γ_{12}			-0.03	(0.14)	-0.09	(0.14)
Science Score	γ_{13}					0.03*	(0.01)
Mathematics Score	γ_{14}					-0.01	(0.01)
Random Effects							
Variance components							
Between-student (τ_{00})		1.753***	(0.17)	1.680***	(0.17)	1.465***	(0.16)
Within-person fluctuation (σ^2)		0.856***	(0.07)	0.859***	(0.07)	0.896***	(0.08)

Notes: no asterisk $p > .05$, *$p < .05$, **$p < .01$, ***$p < .001$; Reference group = Pretest, Control Condition, and average Science and Mathematics Scores

It can be seen from Table 3 that similar to the findings in system organization, there was a significant increase in students' systems behavior scores from pre- to post-test in both conditions (Model 1, $p = .015$). However, when the SES and students' science and mathematics scores were added to the Models, the difference became not significant (Model 3, $p = .095$). Tests of simple slope and difference revealed a significant difference in pretest system behavior scores between students in the computational modeling and paper-based conditions ($t = 2.60$, $p = .01$). The results also showed that a significant increase of system behavior scores was detected for students in the paper-based drawing condition even after SES was controlled and students had average academic scores ($t = 1.94$, two-tailed $p = .05$, one-tailed $p = .027$). In addition, we also found that mathematics score was significantly associated with students' system behavior scores ($p < .001$). Model 3 of system organization accounted for 15% and 23%, respectively, of the within- and between-student variability in systems organization scores. Figure 3 visualizes the results after controlling for students' SES and when students had average academic scores.

System Modeling

For system modeling, the results from a fully unconditional model showed that 66% of the variability in the scores was between-student ($\tau_{00} = 1.49$, $z = 10.00$, $p < .001$), and 34% was within-student ($\sigma^2 = 0.76$, $z = 12.06$, $p < .001$). These results indicated that there was sufficient variability to proceed with MLM. The conditional models were then performed, and the results are presented in Table 4.

The MLM results presented in Table 4 show a different pattern from the results in systems organization and behavior. We found that the increase in system modeling scores was not significant in both conditions before and after SES, science scores, and mathematics scores were added to the models (Model 1 $p = .282$; Model 2 $p = .659$; Model 3 $p = .575$). However,

Model 3 still showed a significant fixed-effect of mathematics scores on students' system modeling scores ($p < .001$). Model 3 accounted for 14% and 20%, respectively, of the within- and between-student variability in systems modeling scores. Figure 3 depicts the results after controlling for students' SES and when students had average academic scores.

Table 4

Unstandardized Coefficients (and Standard Errors) of Multilevel Models of Systems Modeling

Effect	Parameter	Model 1: Condition Effect	Model 2: Controlled SES	Model 3: Controlled SES + Academic Scores
Systems Modeling, β_0				
Intercept	γ_{00}	-0.10 (0.13)	-0.67* (0.29)	-0.23 (0.30)
Condition (Experimental)	γ_{01}	0.63***(0.16)	0.37 (0.21)	0.38 (0.21)
SES	γ_{02}		0.29* (0.14)	0.12 (0.14)
Science Score	γ_{03}			-0.01 (0.01)
Mathematics Score	γ_{04}			0.04***(0.01)
Intervention slope, β_1				
Time (Post)	γ_{10}	0.12 (0.12)	0.12 (0.27)	0.22 (0.28)
Condition (Experimental)	γ_{11}	0.04 (0.15)	0.05 (0.19)	-0.05 (0.20)
SES	γ_{12}		0.00 (0.13)	-0.01 (0.13)
Science Score	γ_{13}			0.03* (0.01)
Mathematics Score	γ_{14}			-0.01 (0.01)
Random Effects				
Variance components				
Between-student (τ_{00})		1.400***(0.14)	1.362***(0.14)	1.194***(0.13)
Within-person fluctuation (σ^2)		0.754***(0.06)	0.754***(0.06)	0.747***(0.06)

Notes: no asterisk $p > .05$, *$p < .05$, **$p < .01$, ***$p <.001$; Reference group = Pretest, Control Condition, and average Science and Mathematics Scores

Overall System Thinking

Last, we ran the same analytical models like the above three systems thinking components on the overall systems thinking scores. The results from a fully unconditional model revealed that 72% of the variability in the overall systems thinking scores was between-student

(τ_{00} =1.36, z = 10.63, p < .001), and 28% was within-student (σ^2 = 0.54, z = 12.05, p < .001). The conditional models were then performed as these results indicated significant variability in the overall systems thinking scores. The results are presented in Table 5.

Table 5

Unstandardized Coefficients (and Standard Errors) of Multilevel Models of Overall Systems Thinking

Effect	Parameter	Model 1: Condition Effect	Model 2: Controlled SES	Model 3: Controlled SES + Academic Scores
Systems Thinking, β_0				
Intercept	γ_{00}	-0.23* (0.12)	-0.91***(0.27)	-0.48 (0.27)
Condition (Experimental)	γ_{01}	0.72***(0.15)	0.41* (0.19)	0.41* (0.19)
SES	γ_{02}		0.34** (0.13)	0.18 (0.13)
Science Score	γ_{03}			-0.00 (0.01)
Mathematics Score	γ_{04}			0.03***(0.01)
Intervention slope, β_1				
Time (Post)	γ_{10}	0.25** (0.10)	0.22 (0.22)	0.34 (0.23)
Condition (Experimental)	γ_{11}	-0.07 (0.12)	-0.07 (0.16)	-0.16 (0.16)
SES	γ_{12}		0.02 (0.11)	-0.01 (0.11)
Science Score	γ_{13}			0.02* (0.01)
Mathematics Score	γ_{14}			-0.00 (0.01)
Random Effects				
Variance components				
Between-student (τ_{00})		1.264***(0.12)	1.215***(0.12)	1.046***(0.11)
Within-person fluctuation (σ^2)		0.520***(0.04)	0.520***(0.04)	0.509***(0.04)

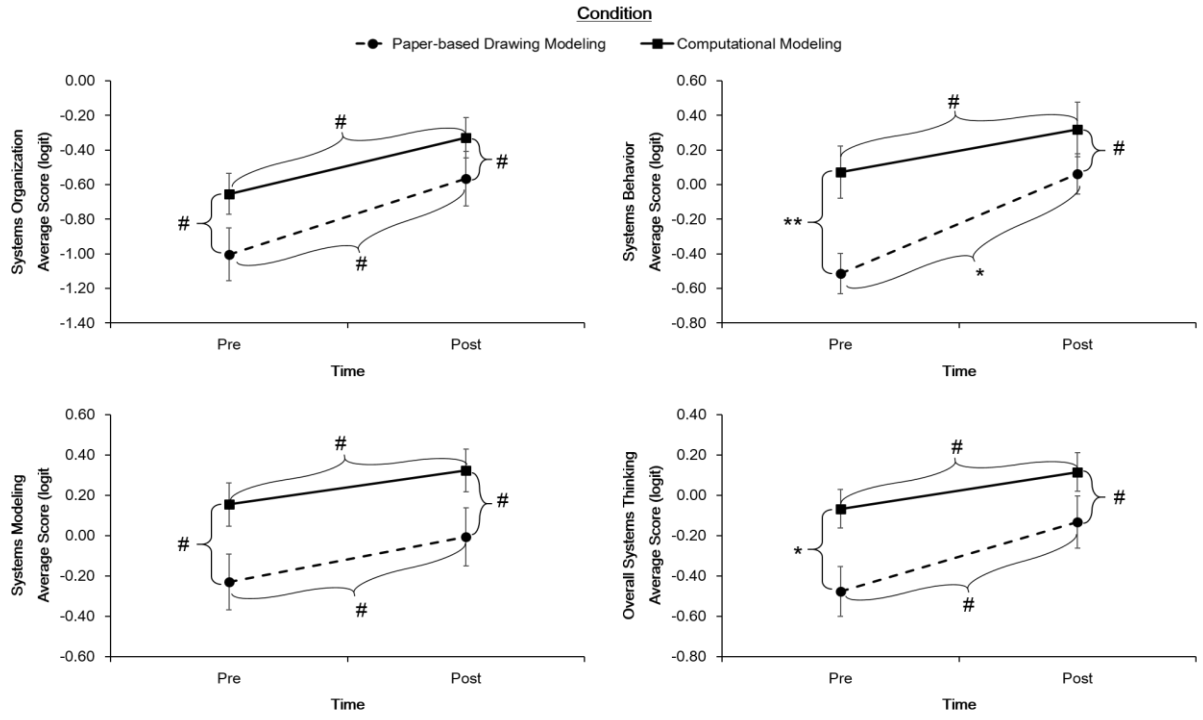
Notes: no asterisk p > .05, *p < .05, **p < .01, ***p <.001; Reference group = Pretest, Control Condition, and average Science and Mathematics Scores

It can be seen from Table 5 that the MLM results indicated a significant increase in overall systems thinking scores from pre- to post-test in both conditions (Model 1, p = .010). Similar to what we found in the systems organization and behavior, this significant increase

became not significant when SES was controlled (Model 2, p = .322) and when students had average academic scores (Model 3, p = .208). Interestingly, even after SES and academic scores were added to the models, the simple slope tests indicated a significant difference in pretest systems thinking scores between those in computational modeling and paper-based modeling conditions (t = 2.21, p = .028). Figure 3 depicts this result. Furthermore, a significant association between overall systems thinking scores and mathematics scores was also found (p < .001). Model 3 accounted for 18% and 23%, respectively, of the within- and between-student variability in overall systems thinking scores.

Figure 3

Score Changes In the Three Components of Systems Thinking and Overall Systems Thinking For Both Conditions Computed Based On Model 3 or After Controlling For SES and Academic Scores



Note: # p > .05, *p < .05, **p < .01

RQ 2. Impacts on Computational Thinking (CT)

A fully unconditional model was run on students' CT scores. The results revealed that 62% of the variability in CT scores was between-student ($\tau_{00} = 0.16$, $z = 9.41$, $p < .001$), and 38% was within-student ($\sigma^2 = 0.10$, $z = 11.67$, $p < .001$). These results indicated that there was sufficient variability in CT scores for subsequent analyses. Therefore, we ran MLM for the three conditional models, and the results are shown in Table 6.

Table 6

Unstandardized Coefficients (and Standard Errors) of Multilevel Models of Computational Thinking

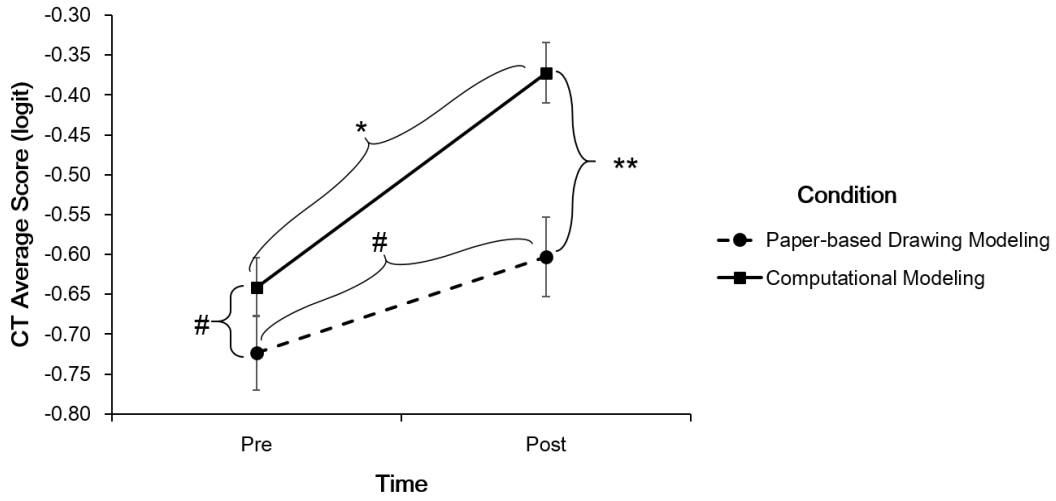
Effect	Para-meter	Model 1: Condition Effect		Model 2: Controlled SES		Model 3: Controlled SES + Academic Scores	
Computational Thinking, β_0							
Intercept	γ_{00}	-0.56***	(0.04)	-0.80***	(0.10)	-0.72***	(0.10)
Condition (Experimental)	γ_{01}	0.22***	(0.05)	0.11	(0.07)	0.08	(0.07)
SES	γ_{02}			0.13**	(0.05)	0.10*	(0.05)
Science Score	γ_{03}					0.01	(0.00)
Mathematics Score	γ_{04}					0.01	(0.00)
Intervention slope, β_1							
Time (Post)	γ_{10}	-0.01	(0.04)	0.07	(0.10)	0.12	(0.10)
Condition (Experimental)	γ_{11}	0.11*	(0.05)	0.15*	(0.07)	0.15*	(0.07)
SES	γ_{12}			-0.04	(0.05)	-0.06	(0.05)
Science Score	γ_{13}					0.00	(0.00)
Mathematics Score	γ_{14}					0.01	(0.00)
Random Effects							
Variance components							
Between-student (τ_{00})		0.145***	(0.02)	0.141***	(0.02)	0.127***	(0.01)
Within-person fluctuation (σ^2)		0.093***	(0.01)	0.095***	(0.02)	0.096***	(0.01)

Notes: no asterisk $p > .05$, *$p < .05$, **$p < .01$, ***$p < .001$; Reference group = Pretest, Control Condition, and average Science and Mathematics Scores

The MLM results for Model 1 showed a significant interaction effect between Time and Condition, indicating a significant difference in the impacts of condition on students' CT scores (p = .035). This result remained significant even after SES was controlled (Model 2, p = .028) and students had average academic scores (Model 3, p = .031). Decomposing Model 3 results, the simple slope tests indicated a significant increase in students' CT scores in the computational modeling condition (t = 1.94, two-tailed p = .053, one-tailed p = .027). However, we did not find this significant result for students in the paper-based pictorial modeling condition (t = 1.19, two-tailed p = .237, one-tailed p = .118). Such an increase for students in the computational modeling condition shifted the CT score from no significant difference between conditions in pretest (t = 1.16, p = .247) to a significant gap of scores in posttest (t = 3.16, p = .002). These results are visualized in Figure 4. Moreover, Model 3 accounted for 13% and 21%, respectively, of the within- and between-student variability in CT scores.

Figure 4

Comparing Changes in CT Scores between the Two Conditions after Controlling For SES and Students Had Average Academic Scores in Model 3



Note: # p > .05, *p < .05, **p < .01

RQ 3. Relationship between CT and Systems Thinking

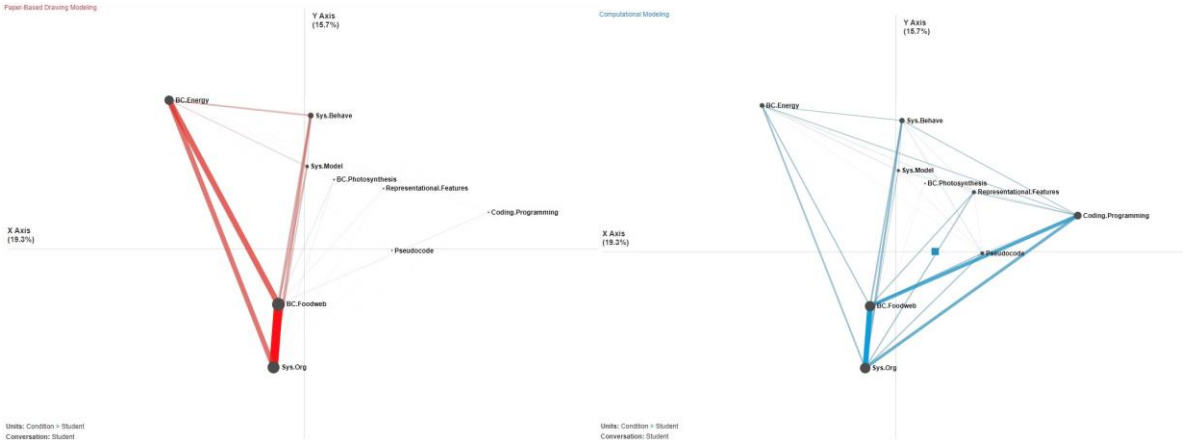
Repeated-measures correlation (r_{rm}) tests were run on CT and the three systems thinking components and the overall systems thinking scores. Based on the results, we found that CT had a weak significant correlation with system organization (r_{rm} = .18, p = .005) and system modeling (r_{rm} = .15, p = .018). In addition, we did not find a significant correlation between CT and system behavior (r_{rm} = .10, p = .129). Moreover, the results also revealed a weak significant correlation between CT and overall systems thinking scores (r_{rm} = .17, p = .008).

RQ 1, RQ 2, and RQ 3. Students' and Teachers' Perceptions: ENA Results for Students' Perceptions

ENA was used to analyze students' responses to an open-ended question in the exit ticket, asking what they thought they learned from each activity. ENA generated a network that visualizes the structure of students' perceptions based upon the qualitative codes we used. Figure 5 shows the structure of students' perceptions in each condition, and Figure 6 depicts students' perceptions in both conditions.

Figure 5

Network Structure of Students' Perceptions in both Conditions

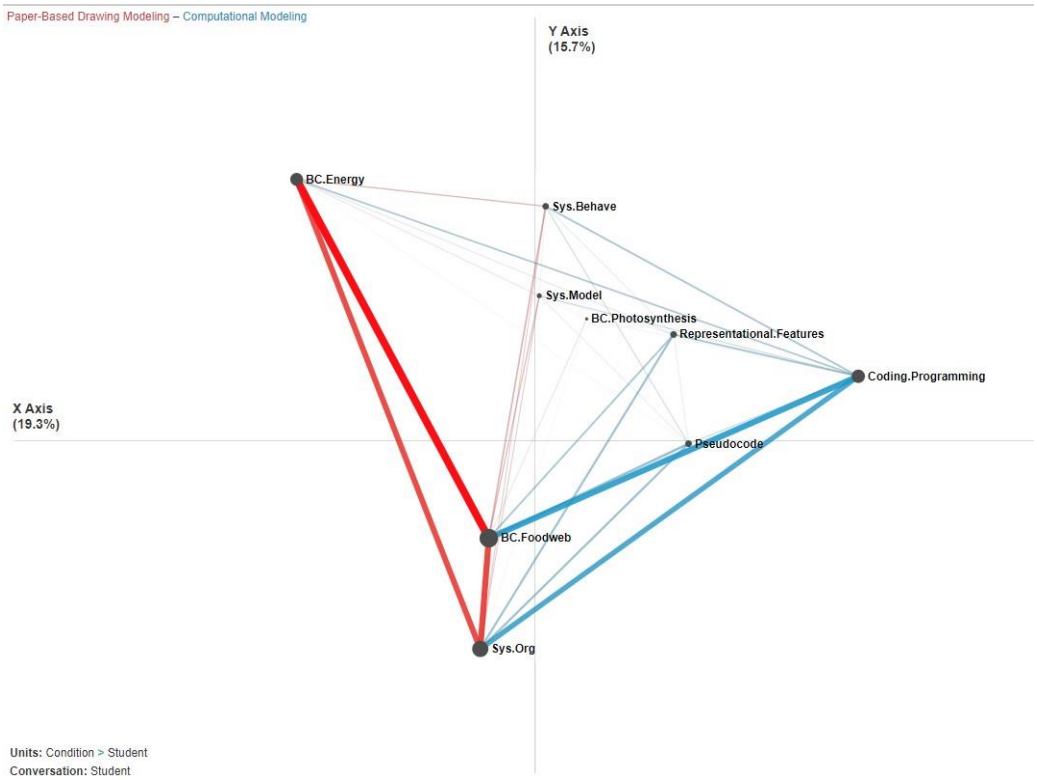


Note: Red Network = Paper-based Modeling Condition; Blue Network = Computational Modeling Condition

It can be seen from Figure 5 that the co-occurrences of codes for students in the paper-based modeling condition, represented in red, were only among biological concepts and systems thinking components. This result indicated that students who were in the paper-based modeling predominantly expressed connections between science concepts. In contrast, there were more co-occurrences of science concepts, systems thinking components, and CS/CT concepts in the computational modeling condition, represented in blue. This demonstrated that students in the computational modeling condition articulated a more integrative understanding of science concepts, systems thinking, and CT.

Figure 6

A Difference Model of Students' Perceptions in both Conditions



Note: Red Network = Paper-based Modeling Condition; Blue Network = Computational Modeling Condition

Further investigation on the network, we found that CS/CT concepts (i.e., coding, programming, and pseudocode) were located along the X-axis. At the same time, systems thinking components were concentrated along the Y-axis. A non-parametric Mann-Whitney *U* test was performed to investigate these ENA findings further. We found that there was a significant difference between paper-based and computational modeling conditions along the X-axis ($U = 26987$, $p < .01$, $r = -.70$) and no significant difference along the Y-axis ($U = 15032$, $p = .41$, $r = .05$). The Mann-Whitney *U* test thus indicated a significant difference in students expressing the CS/CT concepts in the two conditions. In contrast, there was no significant difference of students between conditions expressing systems thinking concepts/components.

Discussion and Implications

RQ 1. Impacts of Modeling Activities on Systems Thinking and Conceptual Understanding of Food Webs

The findings regarding paper-based and computational modeling activities on students' systems thinking skills support Nguyen and Santagata's (2020) findings that both conditions can increase such skills within food web concepts. However, our results also align with studies suggesting the superiority of modeling activities in enhancing students' systems thinking and conceptual understanding (e.g., Bodzin, 2011; Hmelo-Silver et al., 2017; Wilkerson-Jerde & Wilensky, 2015). The current study's findings provide insights into the impacts of the two conditions from an alternative systems-thinking framework developed by Mehren et al. (2018) and Mambrey et al. (2020a), as compared to the element–evidence–causal coherence framework used in Nguyen and Santagata's (2020) study. Analyzing each component in systems thinking skills more closely, we found significant increases only in systems organization and behavior scores. In contrast, no significant increase was detected in students' systems modeling scores.

These findings provide evidence for Mehren et al.'s (2018) and Mambrey et al.'s (2020a) hypotheses and conclusions that the systems organization and behavior components of systems thinking have a shared cognitive ability, while systems modeling is a separate skill from these two constructs. Hence, systems organization and behavior are likely to increase together as the result of a cognitive-focused intervention, whereas such an increase may not be aligned with systems modeling abilities.

An alternative explanation about an insignificant increase in systems modeling scores can be viewed from the learning progression or developmental perspectives. Anderson et al. (2009) argued that middle school students often only grasp the concepts of food webs within a particular system they know or have learned. As the systems modeling skills were assessed by temporal-related questions (where students needed to predict what would happen to the systems in a particular time and condition), the two modeling conditions seemed to lack the support for cognitive development needed to significantly increase such reasoning. Although we included tasks asking students to formulate hypotheses and predict what would happen to agents if they changed some parameters in their codes (e.g., Activity 4 in computational and paper-based modeling activities), these activities did not seem sufficient to allow most students to abstract this skill to other contexts. Further studies are needed to investigate alternate activities or refine the current activities, targeting the frequency of hypothesis generation or prediction-related tasks that would increase students' systems modeling skills.

Regarding the disappearance of significant impacts of the two conditions after controlling for SES, we believe this result might relate to a relationship between SES and the content area within the systems thinking instrument. Many studies have shown a significant positive association between SES and disciplinary understanding (e.g., Baker et al., 2002; Blums et al.,

2017). Such an association might have interfered with the current study's findings resulting in the insignificant impacts of the two conditions on students' systems thinking skills. None of the previous studies exploring the impacts of modeling activities on systems thinking skills included controls for students' SES. Therefore, we believe that this finding corroborates the previous studies related to systems thinking and that researchers should cautiously interpret systems thinking skills assessed with disciplinary content, as SES might affect the results. Furthermore, our results also revealed a significant positive correlation between systems thinking and mathematics scores, indicating, perhaps, a shared cognitive ability. One possible cognitive ability linkage point is spatial ability, which is related to both mathematics scores and computational thinking skills (Román-González et al., 2017; Xie et al., 2020). However, more studies are needed to substantiate this finding.

RQ2. The Impacts of Modeling Activities on CT

The significant positive impact of computational modeling on students' CT skills resonates with findings from Berland & Wilensky (2015) and Aksit & Wiebe (2020). In addition, our finding helps generalize previous studies' findings by providing evidence from a quasi-experimental study and a different scientific concept. In addition, the population came from a culturally distinct country, Indonesia. The computational modeling activities might have fostered the development of students' spatial ability, which is a cognitive ability that underlies CT skills (Román-González et al., 2017). Moreover, the conditional tasks (the use of "if" scenarios) might have increased students' general problem-solving skills whose basis are in fluid intelligence (Green et al., 2017). The reason is that students might cognitively engage in brainstorming different conditions and solutions of a particular decision made to their codes or models to solve specific problems. However, a caveat applies to using the term coding or computer programming

in this study. Although coding, computer programming, and pseudocode are parts of CS or CT, they are neither the main nor the core concepts of CS or CT (Weintrop et al., 2016). However, Vandenberg et al. (2020) found that students at this age are only familiar with these concepts, and thus during the activities and in the analysis (ENA), we used those to represent CT instead of other terms.

Although our findings showed the benefit of computational modeling activities to advance students' CT skills, such modeling activities are still limited to the use of computers (i.e., plugged activities). Because computer availability is integral in the implementation of computational modeling activities, accessibility issues may limit students' exposure to such activities, especially those whose families are from a lower SES. Future studies can develop unplugged (i.e., activities not requiring a computer) activities that focus on CS or CT concepts and practices to increase students' CT skills. Such activities may use the instructional framework developed by Peel et al. (2018, 2019) to be used without computers but can still improve students' understanding of scientific concepts and CT skills. Regardless of this issue, our findings have provided significant insight into how and why computational modeling should be considered an important NGSS 3D instructional approach (NGSS Lead States, 2013). Our study found computational modeling increases disciplinary content, applies crosscutting concepts, and develops science and engineering practices.

RQ3. Relationship between Systems Thinking and CT

We found a significant and weak correlation between systems thinking and CT skills. This result does not align with Weintrop et al.'s (2016) theoretical hypothesis on the relationship between CT and systems thinking, which the authors conjectured that systems thinking is part of CT. Our results revealed that the two constructs were correlated, but they are not likely to bear

on the same skills or reside in the same cognitive dimension. In other words, systems thinking and CT are likely to be separate constructs.

As mentioned earlier, the two constructs may share common cognitive abilities such as spatial ability or even mathematical thinking (Román-González et al., 2017; Xie et al., 2020). However, the relationship between these cognitive abilities and the two constructs is still understudied. Hence, we call for further studies to investigate the association between spatial ability, mathematical thinking, systems thinking, and CT skills. We believe that such studies are crucial to fully understand the underlying science education's important crosscutting concepts and science and engineering practices.

Conclusion, Limitations, and Future Directions

A significant need has emerged for exploring 3D instructional approaches in science that can positively impact students' understanding of disciplinary content, crosscutting concepts, and science and engineering practices. In the current study, we examined the impacts of two modeling activities—paper-based and computational modeling—on seventh-grade students' conceptual understanding of food webs, systems thinking, and CT skills. Our findings revealed that both conditions could increase students' understanding of food web concepts and systems thinking. However, only computational modeling significantly improved students' CT skills. Given that the literature lacks a deep understanding of computational modeling as an instructional strategy, especially regarding its impact on the three NGSS components, the current study bridges the literature gap by providing evidence for the potentially significant impact of computational modeling activities on all three of these components.

Like many studies, the current study has some limitations that may have limited generalization of the results. The first one is the results are limited to the present study's context,

especially in terms of the online delivery mode. Due to the COVID-19 pandemic, students and teachers engaged in online activities that might have been foreign to them, and learning science via computer programming might have accentuated the teachers and students' unfamiliarity with the instructional context. Furthermore, such an online learning environment prevented us from collecting richer data, such as interviews and classroom interactions that could substantiate and corroborate the current quantitative findings. The online environment also limited the generalization of results, given that a majority of students who completed most of the activities were those with above-average academic scores. Therefore, future studies should try to replicate the current studies in a face-to-face setting to gather more evidence for the different impacts of the two conditions on students' science learning. However, it is worth noting that this pandemic-induced online learning did allow for implementation in a foreign country where we might not have otherwise been able to partner with teachers and interact with students.

Second, the current study did not randomly select students or schools for each condition; instead, the selection was purposeful. This approach might have resulted in distribution issues in the systems thinking-embedded food web pretest scores, resulting from uneven ratios of students from higher achieving schools and the quality of teachers. The number of schools and teachers that participated in the current study was relatively small, which prevented us from running more sophisticated quantitative statistical techniques, such as MLM, given such analyses need a larger sample size (> 50; McNeish & Stapleton, 2017). We acknowledge this as one of the current study's limitations, and we encourage future studies to include additional sample schools and teachers to enable them to perform more advanced MLM.

Last, regarding the conceptualization of CT used in the current study, we acknowledge both programming and general problem-solving skills led us to use an instrument with both a

programming context and non-programming context (Wiebe et al., 2019). The questions with a programming context might also relate to the significant increase in students' CT skills as they gained familiarity with block-based programming throughout the activities. Hence, future studies can corroborate our findings by using a programming-free instrument, such as Bebras Challenges only, to assess the efficacy of computational modeling on students' CT skills. However, the researchers should also consider students' testing fatigue, given such an instrument tends to be lengthy. Given that our findings point to the benefits of both plugged and unplugged approaches, future studies are encouraged to develop learning activities with different configurations and combinations of computer and paper-based modeling to enhance students' representational competence.

Acknowledgement

We thank all the students and teachers participating in this study. We also thank Yustika Sya'bandari for being the second coder. This material is based upon the work supported by the National Science Foundation under Grant No. DRL-1640141. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The portions of this study were also supported by NC State University's College of Education and the Friday Institute for Educational Innovation through Dissertation Research Continuity Award and Graduate Student Fellow Award, respectively.

Supporting Materials

Supporting Material A: Psychometric Properties of CTA-M

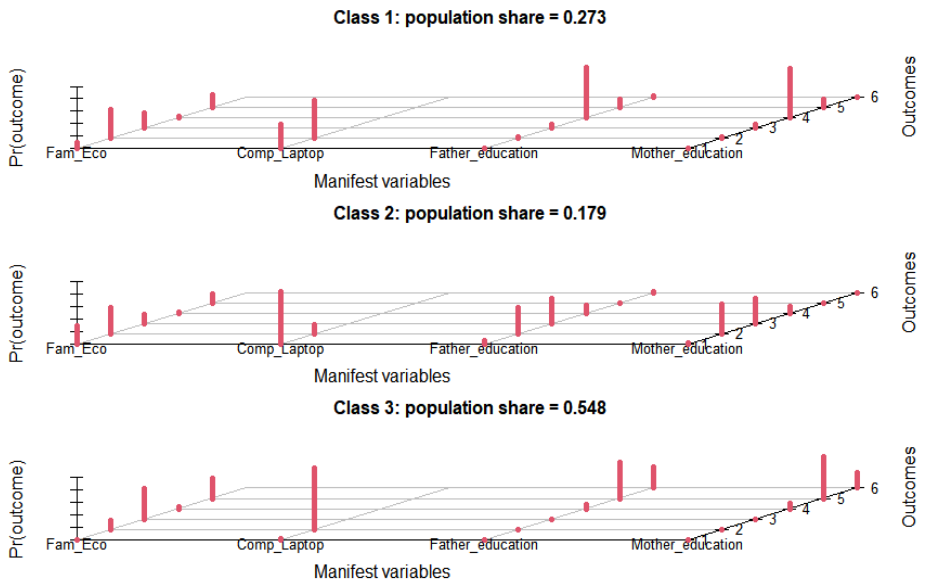
No	Item	25 Items				23 Items			
		Estimate	Unweighted MNSQ	Weighted MNSQ	Alpha if item deleted	Estimate	Unweighted MNSQ	Weighted MNSQ	Alpha if item deleted
1	RG_3	-0.148	0.95	0.95	0.620	-0.033	0.94	0.95	0.657
2	RG_Q7	-0.092	0.93	0.93	0.615	0.024	0.92	0.93	0.653
3	RG_Q11	-0.574	0.96	0.97	0.623	-0.467	0.97	0.97	0.661
4	RG_Q12	0.623	1.11	1.05	0.649	0.749	1.15	1.08	0.684
5	RG_Q13	0.039	1.10	1.09	0.654	0.156	1.15	1.12	0.688
6	RG_Q14	-0.614	0.93	0.94	0.618	-0.508	0.92	0.94	0.655
7	RG_Q16	0.025	0.92	0.92	0.614	0.14	0.92	0.93	0.652
8	RG_Q17	-0.311	1.00	1.00	0.632	-0.2	1.01	1.01	0.669
9	RG_Q18	-0.857	0.94	0.95	0.621	-0.755	0.93	0.94	0.658
10	RG_Q19	0.390	1.13	1.09	0.656	0.512	1.18	1.13	0.69
11	RG_Q20	-0.223	0.95	0.95	0.619	-0.112	0.94	0.95	0.657
12	RG_Q23	1.362	1.31	1.10	0.660				
13	RG_Q24	-0.146	1.05	1.04	0.644	-0.033	1.09	1.07	0.68
14	RG_Q25	0.490	1.06	1.02	0.639	0.614	1.09	1.04	0.675
15	RG_Q26	-0.814	1.05	1.04	0.645	-0.71	1.06	1.06	0.68
16	RG_Q27	-0.772	0.99	1.00	0.632	-0.668	1	1	0.668
17	B_Q1	-0.325	1.00	0.99	0.631	-0.214	1	1	0.667
18	B_Q6	0.437	0.98	0.98	0.629	0.562	1.02	0.99	0.666
19	B_Q7	0.575	1.08	1.05	0.645	0.701	1.11	1.06	0.681
20	B_Q8	0.785	0.99	0.99	0.631	0.915	1.03	1	0.668
21	B_Q9	1.276	1.32	1.11	0.661				
22	B_Q10	0.118	0.98	0.97	0.625	0.237	0.98	0.97	0.662
23	B_Q13	-0.078	1.00	1.00	0.634	0.038	1.02	1.01	0.67
24	B_Q14	-0.619	0.93	0.94	0.619	-0.512	0.92	0.93	0.656
25	B_Q15	-0.546	0.95	0.96	0.622	-0.438	0.95	0.96	0.659
Separation Reliability			.989				.984		
EAP/PV Reliability			.640				.654		
Cronbach's alpha			.643				.678		
Chi-square test of parameter equality			1880.6				1286.21		
df			24				22		
Sig Level			0				0		
Final Deviance			36190.3359				34019.61607		
Akaike Information Criterion (AIC)			36242.3359				34067.61607		
Akaike Information Criterion Corrected (AICc)			36241.24804				34066.69067		
Bayesian Information Criterion (BIC)			36374.01964				34189.17028		
Total number of estimated parameters			26				24		

Supporting Material B: Psychometric Properties of Systems-Thinking-Embedded Food Web Assessment

No	Item	3D 36 Items				3D 32 Items			
		Estimate	Unweighted MNSQ	Weighted MNSQ	α if item deleted	Estimate	Unweighted MNSQ	Weighted MNSQ	α if item deleted
1	SO_1a	-1.801	**1.66**	1.08	0.849				
2	SO_1b	-1.454	0.97	0.97	0.843	-1.472	1.04	0.99	0.827
3	SO_1c	-1.202	1.05	1.01	0.843	-1.219	1.09	1.03	0.827
4	SO_1d	-2.403	1.14	1.04	0.848	-2.415	1.18	1.02	0.834
5	SO_2a	0.339	1.14	1.08	0.837	0.328	1.12	1.07	0.817
6	SO_2b	-0.123	1.14	1.11	0.839	-0.136	1.09	1.1	0.82
7	SO_2c	0.353	1.11	0.99	0.833	0.342	1.02	0.97	0.811
8	SO_2d	0.096	1.01	1.05	0.836	0.082	1.06	1.06	0.816
9	SO_3a	1.8	1.1	0.83	0.829	1.786	1.1	0.85	0.807
10	SO_3b	1.139	0.81	0.85	0.826	1.128	0.83	0.86	0.804
11	SO_3c	1.585	0.74	0.82	0.825	1.577	0.77	0.83	0.803
12	SO_3d	1.671	**1.33**	0.99	0.833				
13	SV_1a	-0.943	1.11	1.03	0.846	-0.757	1.13	1.04	0.827
14	SV_1b	-0.538	1.09	1.02	0.844	-0.336	1.13	1.06	0.825
15	SV_1c	-0.546	0.77	0.89	0.837	-0.343	0.78	0.89	0.814
16	SV_1d	-1.055	1.03	0.95	0.842	-0.867	0.94	0.95	0.821
17	SV_2a	-0.226	0.9	0.94	0.837	-0.008	0.91	0.95	0.815
18	SV_2b	0.579	1.01	0.97	0.837	0.826	1.13	1.04	0.821
19	SV_2c	-0.184	0.87	0.95	0.836	0.036	0.9	0.96	0.814
20	SV_2d	0.553	1.01	0.98	0.839	0.799	1.12	1.04	0.823
21	SV_3a	0.095	0.93	0.98	0.838	0.323	0.99	1.02	0.819
22	SV_3b	1.149	**1.37**	1.1	0.844				
23	SV_3c	0.102	1.04	1.04	0.841	0.328	1.11	1.09	0.822
24	SV_3d	1.014	**1.24**	1.04	0.842				
25	SH_1a	-0.889	1.09	0.98	0.829	-0.903	1.02	0.99	0.829
26	SH_1b	0.05	1.18	1.12	0.831	0.049	1.2	1.11	0.831
27	SH_1c	-0.808	0.92	0.98	0.828	-0.814	0.94	0.99	0.828
28	SH_1d	-0.923	1	0.97	0.828	-0.926	1.01	0.98	0.828
29	SH_2a	0.749	1.19	1.09	0.827	0.764	1.26	1.09	0.827
30	SH_2b	-0.026	0.84	0.92	0.82	-0.018	0.86	0.93	0.82
31	SH_2c	0.651	1.02	0.98	0.822	0.662	1.07	0.99	0.822
32	SH_2d	0.081	0.88	0.93	0.818	0.083	0.86	0.92	0.818
33	SH_3a	0.638	1.06	0.98	0.822	0.641	1.06	0.98	0.822
34	SH_3b	-0.117	0.88	0.97	0.822	-0.123	0.9	0.97	0.822
35	SH_3c	0.575	1.23	1.12	0.828	0.572	1.22	1.12	0.828
36	SH_3d	0.019	0.97	0.99	0.824	0.012	0.97	0.99	0.824
Separation Reliability			.997				.996		
EAP/PV Reliability			.846, .882, .873				.847, .848, .845		
Cronbach's alpha			.849, .852, .837				.832, .835, .837		
Chi-square test of parameter equality			8807.21				6737.11		
df			33				29		
Sig Level			0				0		
Final Deviance			45153.27622				40631.67203		
Akaike Information Criterion (AIC)			45237.27622				40707.67203		
Akaike Information Criterion Corrected (AICc)			45234.50104				40705.39879		
Bayesian Information Criterion (BIC)			45451.05944				40901.09495		
Total number of estimated parameters			42				38		

Supporting Material C: Clustering Analysis for Socioeconomic Statuses

# of Latent Classes	Maximum log-likelihood	AIC	BIC	Likelihood ratio/deviance statistic	X^2 goodness of fit	Note
1	-4447.48	8924.96	8997.14	1723.67	4885.04	-
2	-3846.51	7755.01	7904.19	521.72	642.24	-
3	**-3727.41**	**7548.83**	**7775.01**	**283.54**	**354.98**	-
4	-3706.74	7539.48	7842.66	242.19	314.73	MAXIMUM LIKELIHOOD NOT FOUND
5	-3668.64	7495.28	7875.45	165.99	233.78	MAXIMUM LIKELIHOOD NOT FOUND



REFERENCES

Abrahamson, D., & Sánchez-García, R. (2016). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences*, *25*(2), 203–239. https://doi.org/10.1080/10508406.2016.1143370

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

Aksit, O., & Wiebe, E. N. (2020). Exploring force and motion concepts in middle grades using computational modeling: A classroom intervention study. *Journal of Science Education and Technology*, *29*(1), 65–82. https://doi.org/10.1007/s10956-019-09800-z

Anderson, C. W., Mohan, L., Jin, H., Tan, E., Chen, J., Zesaguli, J., ... & Schramm, J. (2009). Learning progressions for principled accounts of processes in socio-ecological systems. In *Annual Meeting of the National Association for Research in Science Teaching*, Garden Grove, CA.

Anderson, T., Schonborn, K., Du Plessis, L., Gupthar, A., & Hull, T. (2013). Identifying and developing students' ability to reason with concepts and representations in biology. In D. Treagust & C.-Y. Tsui (Eds.), *Multiple Representations in Biological Education* (pp. 19–38). Dordrecht: Springer.

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, *8*, 1–13. https://doi.org/10.3389/fpsyg.2017.00456

Baker, D. P., Goesling, B., & LeTendre, G. K. (2002). Socioeconomic status, school quality, and national economic development: A cross-national analysis of the "Heyneman-Loxley Effect" on mathematics and science achievement. *Comparative Education Review*, *46*(3), 291–312. https://doi.org/10.1086/341159

Batzri, O., Ben-Zvi-Assaraf, O., Cohen, C., & Orion, N. (2015). Understanding the earth systems: Expressions of dynamic and cyclic thinking among university students. *Journal of Science Education and Technology*, *24*(6), 761–775. https://doi.org/10.1007/s10956-015-9562-8

Ben-Zvi-Assaraf, O., & Orion, N. (2005). Development of system thinking skills in the context of earth system education. *Journal of Research in Science Teaching*, *42*(5), 518–560. https://doi.org/10.1002/tea.20061

Ben-Zvi-Assaraf, O., & Orion, N. (2010a). Four case studies, six years later: Developing system thinking skills in junior high school and sustaining them over time. *Journal of Research in Science Teaching*, *47*(10), 1253–1280. https://doi.org/10.1002/tea.20383

Ben-Zvi-Assaraf, O., & Orion, N. (2010b). System thinking skills at the elementary school level. *Journal of Research in Science Teaching*, *47*(5), 540–563. https://doi.org/10.1002/tea.20351

Bergan-Roller, H. E., Galt, N. J., Chizinski, C. J., Helikar, T., & Dauer, J. T. (2018). Simulated computational model lesson improves foundational systems thinking skills and conceptual knowledge in biology students. *BioScience*, *68*(8), 612–621. https://doi.org/10.1093/biosci/biy054

Berland, M., & Wilensky, U. (2015). Comparing virtual and physical robotics environments for supporting complex systems and computational thinking. *Journal of Science Education and Technology*, *24*(5), 628–647. https://doi.org/10.1007/s10956-015-9552-x

Blums, A., Belsky, J., Grimm, K., & Chen, Z. (2017). Building links between early socioeconomic status, cognitive ability, and math and science achievement. *Journal of*

Cognition and Development, *18*(1), 16–40. https://doi.org/10.1080/15248372.2016.1228652

Bodzin, A. M. (2011). The implementation of a geospatial information technology (GIT)-supported land use change curriculum with urban middle school learners to promote spatial thinking. *Journal of Research in Science Teaching*, *48*(3), 281–300. https://doi.org/10.1002/tea.20409

Boersma, K., Waarlo, A. J., & Klaassen, K. (2011). The feasibility of systems thinking in biology education. *Journal of Biological Education*, *45*(4), 190–197. https://doi.org/10.1080/00219266.2011.627139

Bond, T.G., Yan, Z., & Heene, M. (2021). *Applying the Rasch model: Fundamental measurement in the human sciences fourth edition*. New York, NY: Routledge.

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Paper presented at the 2012 Annual Meeting of the American Educational Research Association*. Retrieved from: http://scratched.gse.harvard.edu/ct/files/AERA2012.pdf

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: issues and techniques. *Journal of Advanced Nursing*, *58*(4), 386–395. https://doi.org/10.1111/j.1365-2648.2007.04242.x

Chandrasekharan, S., & Nersessian, N. J. (2015). Building cognition: The construction of computational representations for scientific discovery. *Cognitive Science*, *39*, 1727–1763. https://doi.org/10.1111/cogs.12203

Coll, R. K., & Lajium, D. (2011). Modeling and the future of science learning. In M. Khine & I. Saleh, *Models and modeling* (pp. 3–21). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0449-7_1

Cuny, J., Snyder, L., & Wing, J. M. (2010). *Demystifying computational thinking for non-computer scientists*. Unpublished manuscript in progress, referenced in http:// www.cs.cmu.edu/CompThink/resources/TheLinkWing.pdf.

Denner, J., Werner, L., & Ortiz, E. (2012). Computer games created by middle school girls: Can they be used to measure understanding of computer science concepts?. *Computers & Education*, *58*(1), 240–249. https://doi.org/10.1016/j.compedu.2011.08.006

Deschacht, N., & Goeman, K. (2015). The effect of blended learning on course persistence and performance of adult learners: A difference-in-differences analysis. *Computers & Education*, *87*, 83–89. https://doi.org/10.1016/j.compedu.2015.03.020

Dickes, A. C., Kamarainen, A., Metcalf, S. J., Gün-Yildiz, S., Brennan, K., Grotzer, T., & Dede, C. (2019). Scaffolding ecosystems science practice by blending immersive environments and computational modeling. *British Journal of Educational Technology*, *50*(5), 2181–2202. https://doi.org/10.1111/bjet.12806

DiSessa, A. (1979). On ‘learnable’ representations of knowledge: A meaning for the computational metaphor. In J. Lochhead & J. Clement (Eds.), *Cognitive Process Instruction: Research on Teaching Thinking Skills* (pp. 239-266). Philadelphia, PA: The Franklin Institute Press.

Eberbach, C., Hmelo-Silver, C., Jordan, R., Sinha, S., & Goel, A. (2012, December). Multiple trajectories for understanding ecosystems. In the *Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012)* (Vol. 1, pp. 411–418).

Eilam, B. (2012). System thinking and feeding relations: Learning with a live ecosystem model. *Instructional Science*, *40*(2), 213–239. https://doi.org/10.1007/s11251-011-9175-4

Garcia, D., Harvey, B., & Barnes, T. (2015). The beauty and joy of computing. *ACM Inroads*, *6*(4), 71-79.

Gilbert, J. K. (2008). Visualization: An emergent field of practice and enquiry in science education. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization. Theory and Practice in Science Education* (pp. 3–24). New York, NY: Springer.

Gilbert, J. K., & Treagust, D. (2009). Macro, submicro and symbolic representations and the relationship between them: Key models in chemical education. In E. Gilbert, J. K. Gilbert, & D. Treagust (Eds.), *Multiple Representations in Chemical Education* (pp. 1–8). Dordrecht: Springer.

Green, C. T., Bunge, S. A., Chiongbian, V. B., Barrow, M., & Ferrer, E. (2017). Fluid reasoning predicts future mathematical performance among children and adolescents. *Journal of Experimental Child Psychology*, *157*, 125–143. https://doi.org/10.1016/j.jecp.2016.12.005

Grotzer, T., & Bell, B. (2003). How does grasping the underlying causal structures of the ecosystems impact students' understanding. *Journal of Biological Education*, *38*(1), 16–29. https://doi.org/10.1080/00219266.2003.9655891

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

Hmelo-Silver, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to learn about complex systems. *The Journal of the Learning Sciences*, *9*(3), 247–298.

Hmelo-Silver, C. E., Jordan, R., Eberbach, C., & Sinha, S. (2017). Systems learning with a conceptual representation: a quasi-experimental study. *Instructional Science*, *45*(1), 53–72. https://doi.org/10.1007/s11251-016-9392-y

Hmelo-Silver, Douglas, Holton, & Kolodner. (2009). Designing to learn about complex systems. *Journal of the Learning Sciences*, *9*(3), 247–298. https://doi.org/10.1207/S15327809JLS0903

Hogan, K. (2000). Students' systems reasoning in ecology. *Journal of Biological Education*, *35*(1), 22-28.

Hogan, K., & Fisherkeller, J. (1996). Representing students' thinking about nutrient cycling in ecosystems: Bidimensional coding of a complex topic. *Journal of Research in Science Teaching*, *33*(9), 941–970. https://doi.org/10.1002/(SICI)1098-2736(199611)33:9<941::AID-TEA1>3.0.CO;2-V

Hokayem, H. (2016). Patterns of reasoning about ecological systemic reasoning for early elementary students. *Science Education International*, *27*(1), 117–135.

Hokayem, H., & Gotwals, A. W. (2016). Early elementary students' understanding of complex ecosystems: A learning progression approach. *Journal of Research in Science Teaching*, *53*(10), 1524–1545. https://doi.org/10.1002/tea.21336

Hokayem, H., Ma, J., & Jin, H. (2015). A learning progression for feedback loop reasoning at lower elementary level. *Journal of Biological Education*, *49*(3), 246–260. https://doi.org/10.1080/00219266.2014.943789

Holbert, N. R., & Wilensky, U. (2011). Racing games for exploring kinematics: a computational thinking approach. In *Proceedings of the 7th International Conference on Games +*

Learning + Society Conference (GLS'11) (pp. 109–118). Pittsburgh, PA, USA: ETC Press. Retrieved from http://dl.acm.org/citation.cfm?id=2206376.2206390

Houchins, J. K., Boulden, D. C., & Wiebe, E. (2021, online). Balancing content and computing in scaffolded middle grades computational modeling activities. *Journal of Designs for Learning*.

Hrin, T. N., Milenković, D. D., Segedinac, M. D., & Horvat, S. (2017). Systems thinking in chemistry classroom: The influence of systemic synthesis questions on its development and assessment. *Thinking Skills and Creativity*, *23*, 175–187. https://doi.org/10.1016/j.tsc.2017.01.003

Irgens, G. A, Dabholkar, S., Bain, C., Woods, P., Hall, K., Swanson, H., Horn, M., & Wilensky, U. (2020). Modeling and measuring high school students' computational thinking practices in science. *Journal of Science Education and Technology*, *29*(1), 137–161. https://doi.org/10.1007/s10956-020-09811-1

Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *The Journal of the Learning Sciences*, *15*(1), 11–34. https://doi.org/10.1207/s15327809jls1501_4

Jin, H., Shin, H. J., Hokayem, H., Qureshi, F., & Jenkins, T. (2019). Secondary students' understanding of ecosystems: A learning progression approach. *International Journal of Science and Mathematics Education*, *17*(2), 217–235. https://doi.org/10.1007/s10763-017-9864-9

K-12 Computer Science Framework. (2016). Retrieved from http://www.k12cs.org

Krist, C., Schwarz, C. V., & Reiser, B. J. (2019). Identifying essential epistemic heuristics for guiding mechanistic reasoning in science learning. *Journal of the Learning Sciences*, *28*(2), 160–205. https://doi.org/10.1080/10508406.2018.1510404

Lavi, R., Dori, Y. J., Wengrowicz, N., & Dori, D. (2020). Model-based systems thinking: Assessing engineering student teams. *IEEE Transactions on Education*, *63*(1), 39–47. https://doi.org/10.1109/TE.2019.2948807

Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1996). Children's ideas about ecology 2: Ideas found in children aged 5-16 about cycling of matter. *International Journal of Science Education*, *18*(1), 19–34. https://doi.org/10.1080/0950069960180102

Lee, I., Martin, F., & Apone, K. (2014). Special section education integrating computational thinking across the K–8 curriculum. *ACM Inroads*, *5*(4), 64–71. https://doi.org/10.1145/2684721.2684736

Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., Malyn-Smith, J., & Werner, L. (2011). Computational thinking for youth in practice. *ACM Inroads*, *2*(1), 32–37. https://doi.org/10.1145/1929887.1929902

Lee, T. D., Jones, M. G., & Chesnutt, K. (2019). Teaching systems thinking in the context of the water cycle. *Research in Science Education*, *49*(1), 137–172. https://doi.org/10.1007/s11165-017-9613-7

Lee, Y., Choi, J., & Kim, T. (2013). Discriminating factors between completers of and dropouts from online learning courses. *British Journal of Educational Technology*, *44*(2), 328–337. https://doi.org/10.1111/j.1467-8535.2012.01306.x

Lehrer, R., & Schauble, L. (2012). Seeding evolutionary thinking by engaging children in modeling its foundations. *Science Education*, *96*(4), 701–724. https://doi.org/10.1002/sce.20475

Leonard, J., Buss, A., Gamboa, R., Mitchell, M., Fashola, O. S., Hubert, T., & Almughyirah, S. (2016). Using robotics and game design to enhance children's self-efficacy, stem attitudes, and computational thinking skills. *Journal of Science Education and Technology*, *25*(6), 860–876. https://doi.org/10.1007/s10956-016-9628-2

Linzer, D. A., & Lewis, J. B. (2011). poLCA: An R package for polytomous variable latent class analysis. *Journal of statistical software*, *42*(10), 1-29.

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

Lytle, N., Cateté, V., Boulden, D., Dong, Y., Houchins, J., Milliken, A., ... & Barnes, T. (2019, July). Use, Modify, Create: Comparing computational thinking lesson progressions for STEM classes. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education* (pp. 395–401).

Maani, K. E., & Maharaj, V. (2004). Links between systems thinking and complex decision making. *System Dynamics Review*, *20*(1), 21–48. https://doi.org/10.1002/sdr.281

Mahaffy, P. G., Matlin, S. A., Holme, T. A., & MacKellar, J. (2019). Systems thinking for education about the molecular basis of sustainability. *Nature Sustainability*, *2*(5), 362–370. https://doi.org/10.1038/s41893-019-0285-3

Mambrey, S., Schreiber, N., & Schmiemann, P. (2020b). Young students' reasoning about ecosystems: The role of systems thinking, knowledge, conceptions, and representation. *Research in Science Education*. https://doi.org/10.1007/s11165-020-09917-x

Mambrey, S., Timm, J., Landskron, J. J., & Schmiemann, P. (2020a). The impact of system specifics on systems thinking. *Journal of Research in Science Teaching*, 1–20. https://doi.org/10.1002/tea.21649

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review*, *28*(2), 295–314. https://doi.org/10.1007/s10648-014-9287-x

Mehren, R., Rempfler, A., Buchholz, J., Hartig, J., & Ulrich-Riedhammer, E. M. (2018). System competence modelling: Theoretical foundation and empirical validation of a model involving natural, social and human-environment systems. *Journal of Research in Science Teaching*, *55*(5), 685–711. https://doi.org/10.1002/tea.21436

Meyer, B., Lane, A., & Mullins, J. (2012). *Simulation with Cellular: A project based introduction to programming (first ed.)*. Monash University, Melbourne, Australia. Online: https://github.com/MonashAlexandria/snapapps.

Mohan, L., Chen, J., & Anderson, C. W. (2009). Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching*, *46*(6), 675–698. https://doi.org/10.1002/tea.20314

Moore, T. J., Brophy, S. P., Tank, K. M., Lopez, R. D., Johnston, A. C., Hynes, M. M., & Gajdzik, E. (2020). Multiple representations in computational thinking tasks: A clinical study of second-grade students. *Journal of Science Education and Technology*, *29*(1), 19–34. https://doi.org/10.1007/s10956-020-09812-0

NGSS Lead States. (2013). *Next generation science standards. For states, by states*. Washington, DC: The National Academies Press.

Nguyen, H., & Santagata, R. (2020). Impact of computer modeling on learning and teaching systems thinking. *Journal of Research in Science Teaching*, 1–28. https://doi.org/10.1002/tea.21674

Pallant, A., & Lee, H. S. (2015). Constructing scientific arguments using evidence from dynamic computational climate models. *Journal of Science Education and Technology*, *24*(2–3), 378–395.

Pande, P., & Chandrasekharan, S. (2017). Representational competence: Towards a distributed and embodied cognition account. *Studies in Science Education*, *53*(1), 1–43. https://doi.org/10.1080/03057267.2017.1248627

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York, NY, USA: Basic Books, Inc.

Peel, A., & Friedrichsen, P. (2018). Algorithms, abstractions, and iterations: Teaching computational thinking using protein synthesis translation. *The American Biology Teacher*, *80*(1), 21–28. https://doi.org/10.1525/abt.2018.80.1.21

Peel, A., Sadler, T. D., & Friedrichsen, P. (2019). Learning natural selection through computational thinking: Unplugged design of algorithmic explanations. *Journal of Research in Science Teaching*, *56*(7), 983–1007. https://doi.org/10.1002/tea.21545

Pierson, A. E., & Clark, D. B. (2018). Engaging students in computational modeling: The role of an external audience in shaping conceptual learning, model quality, and classroom discourse. *Science Education*, *102*(6), 1336–1362. https://doi.org/10.1002/sce.21476

Pierson, A. E., Brady, C. E., & Clark, D. B. (2020). Balancing the environment: Computational models as interactive participants in a STEM classroom. *Journal of Science Education and Technology*, *29*(1), 101–119.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Reiner, M., & Eilam, B. (2001). Conceptual classroom environment-a system view of learning. *International Journal of Science Education*, *23*(6), 551–568. https://doi.org/10.1080/095006901300172458

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

Royal Society. (2012). *Shut down or restart: The way forward for computing in UK schools*. Retrieved from https://royalsociety.org/~/media/education/computing-in-schools/2012-01-12-computing-in-schools.pdf

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, *92*(3), 499–525. https://doi.org/10.1002/sce.20264

SAS Institute. (1997). *SAS/STAT software: Changes and enhancements through release 6.12*. Cary: SAS Institute.

Scherer, H. H., Holder, L., & Herbert, B. (2017). Student learning of complex earth systems: Conceptual frameworks of earth systems and instructional design. *Journal of Geoscience Education*, *65*(4), 473–489. https://doi.org/10.5408/16-208.1

Scherr, R. E., & Robertson, A. D. (2015). Productivity of "collisions generate heat" for reconciling an energy model with mechanistic reasoning: A case study. *Physical Review Special Topics-Physics Education Research*, *11*(1), 1–16. https://doi.org/10.1103/PhysRevSTPER.11.010111

Schreier, M. (2012). *Qualitative content analysis in practice*. London, UK: Sage Publishing.

Selby, C., & Wollard, J. (2013). Computational thinking: The developing definition. In *ITiCSE Conference 2013* (pp. 5–8). ACM.

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

Shaffer, D. W., & Ruis, A. R. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 175–187) Society for Learning Analytics Research.

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

Siebert-Evenstone, A., & Shaffer, D. W. (2019, October). Cause and Because: Using Epistemic Network Analysis to Model Causality in the Next Generation Science Standards. In B. Eagan, M. Misfeldt, A. Siebert-Evenstone, *International Conference on Quantitative Ethnography* (pp. 223–233). Springer, Cham.

Stratford, S. J., Krajcik, J., & Soloway, E. (1998). Secondary students' dynamic modeling processes: Analyzing, reasoning about, synthesizing, and testing models of stream ecosystems. *Journal of Science Education and Technology*, *7*(3), 215–234. https://doi.org/10.1023/A:1021840407112

Suarez, M. (2003). Scientific representation: Against similarity and isomorphism. *International Studies in the Philosophy of Science*, *17*(3), 225–244. https://doi.org/10.1080/0269859032000169442

Sweeney, L. B., & Sterman, J. D. (2007). Thinking about systems: student and teacher conceptions of natural and social systems. *System Dynamics Review: The Journal of the System Dynamics Society*, *23*(2-3), 285–311. https://doi.org/10.1002/sdr.366

Tang, X., Yin, Y., Lin, Q., Hadad, R., & Zhai, X. (2020). Assessing computational thinking: A systematic review of empirical studies. *Computers & Education*, *148*(January), 1–22. https://doi.org/10.1016/j.compedu.2019.103798

Vandenberg, J., Tsan, J., Boulden, D., Zakaria, Z., Lynch, C., Boyer, K. E., & Wiebe, E. (2020). Elementary students' understanding of CS terms. *ACM Transactions on Computing Education (TOCE)*, *20*(3), 1–19. https://doi.org/10.1145/3386364

Von Bertalanffy, L. (1968). *General system theory: Foundations, development, applications*. New York, NY: Geroge Braziller.

Wade-Jaimes, K., Demir, K., & Qureshi, A. (2018). Modeling strategies enhanced by metacognitive tools in high school physics to support student conceptual trajectories and understanding of electricity. *Science Education*, *102*(4), 711–743. https://doi.org/10.1002/sce.21444

Wagh, A., Cook-Whitt, K., & Wilensky, U. (2017). Bridging inquiry-based science and constructionism: Exploring the alignment between students tinkering with code of computational models and goals of inquiry. *Journal of Research in Science Teaching*, *54*(5), 615–641. https://doi.org/10.1002/tea.21379

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

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

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, *8*(1), 3–19. https://doi.org/10.1023/A:1009421303064

Wilkerson-Jerde, M. H., & Wilensky, U. J. (2015). Patterns, probabilities, and people: Making sense of quantitative change in complex systems. *Journal of the Learning Sciences*, *24*(2), 204–251. https://doi.org/10.1080/10508406.2014.976647

Wing, J. M. (2008). Computational thinking and thinking about computing. P*hilosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *366*(1881), 3717–3725. https://doi.org/10.1098/rsta.2008.0118

Xie, F., Zhang, L., Chen, X., & Xin, Z. (2020). Is spatial ability related to mathematical ability: A meta-analysis. *Educational Psychology Review*, *32*, 113–155. https://doi.org/10.1007/s10648-019-09496-y

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 (TOCE)*, *14*(1), 1–16.

Yin, Y., Hadad, R., Tang, X., & Lin, Q. (2020). Improving and assessing computational thinking in maker activities: The integration with physics and engineering learning. *Journal of Science Education and Technology*, *29*(2), 189–214. https://doi.org/10.1007/s10956-019-09794-8

You, H. S., Marshall, J. A., & Delgado, C. (2018). Assessing students' disciplinary and interdisciplinary understanding of global carbon cycling. *Journal of Research in Science Teaching*, *55*(3), 377–398. https://doi.org/10.1002/tea.21423

CHAPTER 3: Exploring Middle School Students' Interests in Computationally Intensive Science Careers: The CISCI Instrument Validation and Intervention

Abstract

Computationally intensive science-related disciplines and careers are predicted to be among the cutting-edge fields and high-demand jobs. Such disciplines and careers demand expertise in both traditional scientific knowledge and computer science–related understanding and skills. Preparing educational systems for developing interest and abilities for such career paths requires a new set of research tools. In this study, we developed and validated a multidimensional instrument called CISCI to measure middle school students' interests in computationally intensive science careers. We also explored several predictors of students' interests in such careers and examined the impact of a scientific computational modeling activity on students' interests. A total of 934 middle school students (aged 11–14) participated in this study. A combination of the classical test theory (CTT) and item response theory (IRT) approaches was used to validate the CISCI instrument. Multiple linear regression tests were run to identify the predictors of students' career interests, and paired-sample *t*-tests were used to examine a scientific computational modeling activity's impact on students' career interests. The results revealed that CISCI was a psychometrically valid and reliable instrument to measure students' career interests. We also found that science and CS attitudes, computational thinking, and prior experience in CS-related activities were significant predictors of students' career interests. In addition, computational modeling activity significantly influenced the frequency of students discussing such career paths with their parents. This study underscores the importance of engaging students in CS-integrated science learning activities to help develop their interests in computationally intensive science jobs.

Keywords: career interest, computational science, intervention, middle grades, validation

Introduction

Recent decades have seen a growing number of career paths in science, technology, engineering, and mathematics (STEM) fields. The proliferation of such career opportunities has been challenging for many countries as they attempt to recruit and prepare students in a STEM pipeline (Malcom & Feder, 2016; Schwab & Sala-i-Martín, 2012). In Europe, there were 6.6 million STEM jobs in 2015, with this number predicted to rise by 6.5% by 2025 (European Center for the Development of Vocational Training, 2014; Shapiro et al., 2015). In the same year, there were 8.6 million STEM jobs in the United States (U.S. Bureau of Labor Statistics, 2017). Of these, 45% required employees to be well trained in computer skills, a percentage which is projected to increase by 12.5% between 2014 and 2024 (U.S. Bureau of Labor Statistics, 2017). This accelerated growth of computationally-intensive STEM career pathways has pushed policymakers to place greater emphasis on K–12 CS education (National Academies of Sciences, Engineering, and Medicine, 2018). Current offerings for CS courses and experiences are uneven within and across states due to the lack of uniform requirements for CS-related coursework and curricula at the state, national, and even global levels (Barr & Stephenson, 2011; Grover & Pea, 2018; Loyalka et al., 2019).

Presently, integrating CS and computational thinking (CT)–related experiences into other STEM courses represents a popular strategy to increase exposure to computing in K–12 education (Goode & Chapman, 2016). For instance, in the U.S., Next Generation Science Standards (NGSS) have listed CT as one among eight essential science and engineering practices (NGSS Lead States, 2013), helping support the proliferation of computationally-intensive curricula and providing powerful justification for CS and CT being part of science and other

STEM coursework. CT is defined as a set of abilities that facilitate abstract and algorithmic thinking, often by using CS tools and concepts to solve problems (Grover & Pea, 2018; Shute et al., 2017). In science learning activities, students engage in CT mostly through computationally rich science activities (Weintrop et al., 2016). Hsu et al. (2018) noted that the most popular computationally rich science activities involved students working to build scientific computational models using computer programming. Research has suggested that computationally rich activities have been found to improve students' understanding of scientific concepts (e.g., Aksit & Wiebe, 2020; Peel et al., 2019); scientific reasoning (e.g., Dickes et al., 2016; Irgens et al., 2020; Pallant & Lee, 2015; Pierson & Clark, 2018); representational fluency (Moore et al., 2020); science interest and motivation (e.g., Wagh et al., 2017); and epistemic agency (Irgens et al., 2020; Pierson et al., 2020).

Of particular interest in this study is how students' learning experiences with computationally rich science activities may support the development of their interests in computationally-intensive science careers (Navlakha & Bar-Joseph, 2011). According to National Academies of Sciences, Engineering, and Medicine (NASEM, 2018) these particular careers areas are growing and "require expertise in the traditional domain and a general fluency in tools and methods from computer science" (p. 3).

Lent et al.'s (1994) socio-cognitive career theory (SCCT) is a framework well-suited to help explain how learning experiences influences career interests, including career and academic self-efficacies and values. Such a theoretical tool can help measure the effectiveness of academic interventions designed to increase the number of students interested in and people entering computationally intensive science (CIS) jobs (Falloon et al., 2020; Goode & Chapman, 2016). However, few studies have leveraged this theoretical framework to specifically look at this issue,

which Bortz et al. (2019) indicated has resulted in a lack of psychometrically sound instruments to measure both students' cognitive and affective-motivational dimensions that are related to CIS career pathways. This study, therefore, aims to develop such an instrument, based on SCCT and targeting the context of CIS careers.

Middle grade students (those aged 11–14) are the focus of this study since many of the computationally rich science activities that have been developed are intended for this age range (e.g., Aksit & Wiebe, 2020; Pallant & Lee, 2015; Wilkerson-Jerde et al., 2015). In addition, this age range has been identified as a critical juncture, when students start to crystalize their interests and identify which careers they want to pursue in the future (Kier & Blanchard, 2020; Kier et al., 2014; Skamp, 2007; Wiebe et al., 2018). While many studies have developed instruments to measure students' interests in STEM careers (e.g., Kier et al., 2014; Roller et al., 2020; Shin et al., 2018), they do not specifically target computationally-intensive careers and thus cannot be used to inform whether particular computational-related activities influence students' interests in such career paths. Therefore, the current study aimed to develop and validate such an instrument. In addition to instrument development, the present study examined the predictors of students' interests in computationally intensive science careers by using their demographics, affective, and cognitive characteristics. Moreover, the effectiveness of the instrument was also evaluated through an intervention of computationally rich science activities. A theoretical framework and other related work are provided in the following sections.

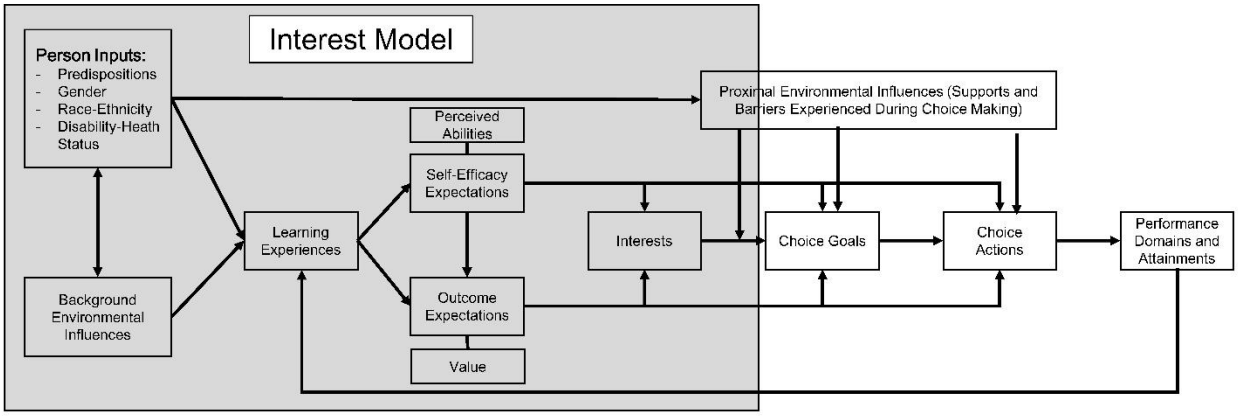
Theoretical Framework: Social Cognitive Career Theory (SCCT) Interest Model

SCCT is based on Bandura's (1986) social cognitive theory of learning, which explains the importance of self-efficacy and outcome expectations in one's actions (Lent et al., 1994, 2000). Self-efficacy is defined as an individuals' beliefs and confidence in their abilities "to

organize and execute courses of action required to attain designated types of performances" (Bandura, 1986, p. 391). Outcome expectations are "personal beliefs about the consequences or outcomes of performing particular behaviors" (Lent et al., 2002). Outcome expectations are also shaped by career/work and academic values as these factors usually set individuals' preferences for work or career conditions (Lent et al., 2002). Lent et al. (2002) noted that SCCT also incorporates a third important component, called personal goals—determination to engage in a particular event to influence certain future outcomes. These three constructs interact together and build three different models of SCCT, namely interest, choice, and performance (Lent et al., 2002).

Figure 1

Social Cognitive Career Theory's Interest Model Adopted from Lent et al. (1994, 2002)



Our current study was framed by SCCT's interest model depicted in Figure 1. SCCT's interest model stresses the cognitive and experiential factors that engender career-related interests, "while tracing the role of interests in helping to motivate choice behavior and skill acquisition" (Lent et al., 2002, p. 265). Moreover, as visualized in Figure 1, this model emphasizes self-efficacy and outcome expectations as significant direct influencers on interests.

Lent et al. (2000) also recognized the influence of contextual factors as a part of this model. These factors included family support and role models, as well as learning experiences—not as a direct influence on students' interests, but indirectly through self-efficacy and outcome expectations. Furthermore, SCCT is also concerned with the ultimate impact of these factors– learning experiences, self-efficacy, outcome expectations, and interests – on career choices and aspirations.

The SCCT interest model, together with other previous instrument development and validation studies such as by Kier et al. (2014), Roller et al. (2020), and Shin et al. (2018), framed and inspired the generation of five constructs to measure students' computationally intensive science careers or CISCI (heretofore CISCI) instrument. These five constructs are: (1) CIS academic self-efficacy, (2) CIS career self-efficacy, (3) CIS career value, (4) CIS career interests and goals, and (5) CIS parental supports and role models. *Academic self-efficacy* is defined as students' beliefs about their ability to successfully achieve a specific academic goal. In the context of this study, this construct measures students' confidence in whether they can learn and develop scientific knowledge through computer programming. *Career self-efficacy* refers to individuals' beliefs about their ability to successfully achieve a specific career goal, which in this study refers to a CIS career. *Career value* measures students' beliefs about whether a particular career is personally meaningful and purposeful to them. In this study, it is conceptualized as whether CIS careers are important to students. The construct of *Career interests and goals* deals with students' interests in pursuing a CIS career. Last, the construct of *parental supports and role models* measures the contextual factors pertaining to the development of students' interests in pursuing CIS careers. This construct measures whether students are

actively discussing such careers with their parents or guardians, and whether they have a role model they can reflect on or consult with in their immediate and extended families.

Hypothetical Predictors of Interest in Computationally Intensive Science Careers

SCCT explains factors that significantly influence one's interests and career development goals: self-efficacy, outcome expectations, and contextual or personal inputs or demographics factors (Lent et al., 1994, 2002). NASEM (2018) stated that CIS careers require both expertise in traditional scientific disciplines and CS. Therefore, we approached the hypothetical predictors by reviewing influencing factors of students' interests in both science and CS careers, as well as general STEM careers.

Cognitive Performance

Personal performance accomplishments are among the most influential factors for self-efficacy and outcome expectations (Bandura, 1986; Lent et al., 1986). In science, many researchers have found that science academic grades or previous performance in science-related activities predict students' interests in science-related careers (e.g., Dabney et al., 2012; Hazari et al., 2013; O'Brien et al., 1999). For instance, Hazari et al. (2013) explored the predictors of female students' interests in physical science–related careers and found that their science grades significantly influence their interests. In addition, the authors also found a significant impact of mathematics grades on students' career interests. Through studying the impact of out-of-school activities on students' career interests, Dabney et al. (2012) also found that mathematics achievement significantly predicts students' interests in STEM careers. Likewise, a similar pattern has been identified in research on CS-related career interests. Smith (2002) conducted a correlational study using path analysis to test a model of the impact of performance on students' interests in information technology careers and majors. The author found that past performance

significantly influences interest mediated by self-efficacy and outcome expectations. Consistent with this finding, Alshahrani et al. (2018) conducted a qualitative study looking at the motivation of students' choices and interests in CS majors, where they found that students' previous achievement played a crucial role. A study by Sáinz and Eccles (2012) found that students' interests in CS and computing significantly depended on their mathematics ability. Using a different approach through SEM, Hava and Ünlü (2021) found a positive relationship between computational thinking dispositions (i.e., students' tendencies to use critical, cooperative, algorithmic, and creative thinking) and career interests in mathematics, as well as other STEM subjects. However, this study focused on affective dimensions and did not consider CT as a cognitive factor. Given how it was measured, the study has provided a piece of evidence for the reciprocal relationships between cognitive and affective constructs among science, mathematics, and CS/CT. Moreover, it also motivated us to investigate more deeply the association between career interests and CT when it is measured cognitively.

Attitudes toward Academic Subjects

Several studies have found that performance in certain academic subjects interact with attitudes toward these subjects, influencing career interests (e.g., Ong et al., 2020; Price et al., 2019; Wiebe et al., 2018). For example, Price et al. (2019) found that science attitudes significantly correlated with STEM career interests. Wiebe et al. (2018) surveyed 11,850 elementary- to high school-level students to determine their attitudes and career interests regarding STEM disciplines and found significant correlations between STEM attitudes and STEM career interests. Similarly, in CS, Funke et al. (2016) found that students' perceptions of attitudes toward CS and computing were significant factors in deciding to major in CS.

Prior Experience

According to Bandura (1997), past experiences, especially successful ones, positively shapes one's self-efficacy and outcome expectations. In CS, Hinckle et al. (2020) confirmed this view by performing path analysis and found a significant influence of prior participation in CS-related activities on CS attitudes. For science, Kong et al. (2014) researched the impact of science summer camp experiences on middle school students' interest in science and engineering careers. They found that this experience increased student interest in such careers. A more recent study conducted by Jones et al. (2021) using SEM also found that prior STEM experiences significantly predicted students' science career aspirations. However, mixed findings exist in the literature regarding the impact of computationally rich science activities. Leonard et al. (2016) conducted a pilot study on the impact of robotics activities and game design on eighth-grade students' CT skills and STEM attitudes and careers. They found that students who were involved in game-design activity exhibit higher CT skills after the activity compared to those in a robotics activity. However, the authors found that neither students in robotics nor game-design activities significantly increased their STEM attitudes and career interests. These mixed findings leave open for discussion the relationship between CT, CS, and STEM attitudes and career interests. We hypothesized that these mixed findings were due to the lack of specificity in their career interest measures, which were probably too general and neglected the other contextual issues, such as students' attitudes toward each STEM discipline, their academic performance, and their demographic backgrounds (e.g., SES and gender).

Demographic Factors

Research has widely indicated that students' gender, grade, and family socioeconomic (SES) level significantly predict their science, STEM, and CS attitudes and career interests (e.g.,

Rachmatullah et al., 2018; Saw, Chang, and Chan, 2018; Turner et al., 2019; Wiebe et al., 2018). For instance, in both science and CS (and more broadly in STEM), male students have been consistently found to have more positive attitudes toward and interest in pursuing careers related to those fields than female students (e.g., Hinckle et al., 2020; Rachmatullah et al., 2018). This gender disparity might be due to the prevalent historical marginalization of females in both fields (NASEM, 2020). Regarding grade level, the literature has shown that many students' interests and attitudes decline over time in science and CS subjects (e.g., Osborne et al., 2003; Shin et al., 2018). A study by Shin et al. (2018) on Korean and Indonesian middle and high school students' STEM career interests found the same declining pattern of STEM career interest in both countries as students moved from middle to high school. The third significant predictor of career interest is the SES level. Turner et al. (2019), using SEM, found that SES did not significantly influence STEM self-efficacy; however, it had a significant positive relationship on students' STEM outcome expectations. Saw et al. (2018) corroborated Turner et al.'s findings by conducting a cross-sectional and longitudinal study to examine the impact of SES on STEM career interests directly. The authors found that students from lower SES backgrounds tended to have lower STEM career interests than students from higher SES backgrounds. They also found that this SES-career interest relationship interacted with gender and ethnicity, suggesting underrepresented minority groups tended to have low STEM career interests.

Research's Purposes and Questions

The current study's purpose was twofold. The first purpose was to validate an instrument to measure middle students' interests in computationally rich science careers, called the Computationally Intensive Science Career Interests (CISCI) instrument. Secondly, this study aimed at examining the significant predictors of interest in such careers by accounting for

cognitive, affective, and demographic factors. The following research questions guided this study:

1. Is the CISCI instrument valid and reliable?
2. What are the significant predictors of students' interests in computationally intensive science careers?
3. What is the impact of participating in computationally rich science activities on students' interest in computationally intensive science careers?

Methods

Participants and Context

The participants in this study were 934 Indonesian middle school students (grades 7-9, age M =12.85) from eight middle schools located in western Java Island and southern Sumatra Island. Regarding gender, 50% of the participating students were females, 48% were males, and the remaining 2% did not provide information about their gender. Almost half of the students' self-reported high socioeconomic (SES) status (55%), 27% from middle status families, and 18% were categorized as from low status. Moreover, in terms of their prior experience in computer science-related activities (e.g., computer programming/coding), 24% had some experiences, while the remaining students (72%) indicated low experience in such activities.

Of 934 students (Full Dataset), 224 seventh-grade students participated in a week-long computationally rich science activity. In this activity, students were engaged in building a computational model representing food webs via block-based programming (more details about the activity see Lytle et al., 2019 and Rachmatullah et al., 2021). This activity was intended to examine how the validated CISCI instrument measured students' response to a computationally-intensive intervention. We only used data from students who completed all the tasks associated

with the activity, thus only data from 186 of the 224 who participated in the intervention (Intervention Data-subset) were used in the further analyses to answer the third research question. The Full Dataset was used to validate the CISCI instrument and explore the predictors of students' CIS career interests.

Instrument Development and Validation Procedure: Computationally Intensive Science Career Interests (CISCI) Instrument

CISCI instrument consisted of five constructs based on SCCT: academic self-efficacy, career self-efficacy, career value, career interest and goal, and parental and contextual support. The items were in the Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). All questions were adapted from Kier et al. (2014) and Shin et al. (2018) STEM career measures. Table 1 shows all the initial items before removal.

The validation procedure of CISCI instrument was guided by Messick's (1995) construct validity process and the Standards for Educational and Psychological Testing proposed by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (AERA, APA, & NCME, 2014). AERA et al. (2014) suggest that researchers need to address five sources of validity evidence when they validate an instrument: response processes, test content, consequences of testing, internal structure and relation to other variables or criterion validity. Nonetheless, Messick (1995) argues that validation is an ongoing and iterative process and researchers may start by focusing on one or two sources of validity evidence. In this study, we focused on the last four sources.

Content Validity

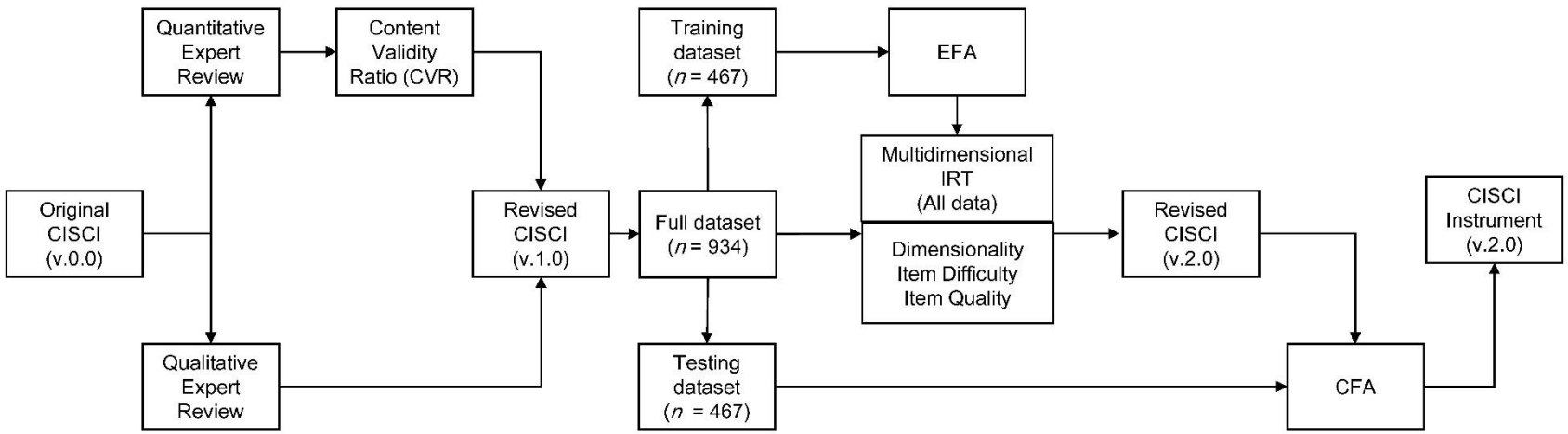
AERA et al. (2014) define test content as "an analysis of the relationship between the content of a test and the construct it is intended to measure" (p. 14). This source of validity

evidence concerns the wording and themes of the items, and can be addressed through expert review. Hence, after a set of questions for each construct was created, three experts in education were asked to qualitatively evaluate the items. These experts focused on the wording, readability, and appropriateness of the items (i.e., whether they measure the constructs or not). After the experts provided feedback on the instrument, the instrument was revised.

The revised instrument was then sent to 11 other experts in middle-grades education, science education, and STEM education. The experts were professors, teachers, graduate students, and researchers. This process was guided by Lawshe's (1975) content validity ratio. The experts were asked to provide feedback on each item by selecting either "essential" or "useful but not essential" or "not necessary." Their responses were utilized to calculate a content validity ratio (CVR) for each item using the following formula: CVR = $(n_e - N/2)/(N/2)$ where n_e refers to the total number of experts who indicate that it is "essential" to measure the degree to which students are interested in computationally demanding science careers. N refers to the total number of experts. The CVR values can range between -1.0 to +1.0, and only the items that meet the cutoff value are retained. Ayre and Scally (2014) suggested that the cutoff value with 11 experts is 0.818. We did not remove the items that were below the cutoff; instead, we revised them by considering suggestions from the expert to maintain having three items per construct (MacCallum et al., 1999; Raubenheimer, 2004). We then calculated the content validity index (CVI) which is the average CVR, and used the same cutoff to interpret the whole instrument's content validity.

Figure 2

The Flow of Validation Procedure for CISCI Instrument



The revised instrument was then administered to students and used to begin addressing the internal structure validity. AERA et al. (2014) define internal structure validity as "the degree to which the relationships among test items and test components conform to the construct on which the proposed test score interpretations are based." This definition reflects Messick's (1995) definition of construct validity. Hence, validity related to the structure of the instrument: the number of latent constructs (factors and dimensions are used interchangeably with constructs throughout the paper), item difficulty, and item quality are part of obtaining evidence of internal structure validity and of this study's area of interest. To do so, this study combines classical test theory (CTT) and item response theory (IRT) approaches to address the internal structure validity. The data from the full dataset of 934 students was divided equally into two datasets, one dataset (training set) for Exploratory Factor Analysis (EFA) and another dataset (testing set) was used in Confirmatory Factor Analysis (CFA). For IRT-Rasch, we used the full dataset so that we could gather more evidence for the quality of each item using analytic techniques not available in EFA and CFA. Figure 2 visualizes the flow of the validation procedure.

Construct Validity

Exploratory Factor Analysis (EFA). According to Hair et al. (2019), EFA's main purpose is "to define the underlying structure among the variables in the analysis" (p. 124). The number of underlying factors generated by EFA is based not on the theoretical assumptions, but on the data. The EFA results, based on data, can then be compared to a factor structure predicted by theory and prior empirical findings. Hair et al. (2019) suggested that researchers need to consider the testing assumptions before running EFA on the training dataset. The three assumptions include the conceptual foundation that assumes the existence of factors (we used factor, dimension, and construct interchangeably in this paper) in the instrument, the existence of

sufficient correlation among variables shown by a significant result ($p < .05$) in Bartlett's test of sphericity, and the measure of sampling adequacy shown by a Kaiser-Meyer-Olkin test exceeds .50. After meeting all these assumptions, the EFA was run on the dataset. Both parallel analysis and eigenvalues were used to determine the number of factors in the instrument (Hair et al., 2019; Schmitt, 2011). Scree plots were also used to consider the number of factors. For the rotation methods, Varimax was used in this study, given that it shows simpler separation of the factors than other rotation methods (Hair et al., 2019). The assumption tests were run in IBM SPSS version 26 (IBM Corp, 2018), and EFA was run using the **fa** function from the **psych** package in R (Revelle, 2018).

Multidimensional Item Response Theory (IRT). Multidimensional IRT was used to evaluate the structure and quality of the instrument as well as the difficulty of the items in the instrument. In addition to theory, the multidimensional analysis can both be informed by the EFA results and aid in the interpretation of the EFA findings. This analysis enabled us to identify the best model and misfitting items within each dimension (Boone, 2016). Misfitting items here refer to the psychometric behavior of the items based on the patterns of students' responses to the items. This analysis also considered how well the items could differentiate high and low interest students. Boone et al. (2014) suggested using mean-square (MNSQ) values to identify misfitting items or the quality of the items. Wright and Linacre (1994) indicated that a well-behaved item has an MNSQ value between 0.60 and 1.40, and this cutoff was used to evaluate the items in this study.

The model is considered improved when it has fewer misfitting items than the competing models. In this study, four different models were compared. The first model was that of one-dimension CISCI, assuming that the instrument is one dimension (baseline), the second and their

models were suggested by the EFA results based upon parallel analysis and eigenvalues, and the fourth model was the five-dimension model based on the theoretical framework. Adams and Wu (2010) suggested identifying the best model by looking at the lowest chi-square, final deviance, and Akaike information criterion (AIC). If the best model was identified, but still consisted of misfitting items, then those items were removed, and this process was run in several iterations until no misfitting items were found. Multidimensional IRT was run in ConQuest version 5.12.3 (Adams et al., 2020). CFA was run after both EFA and multidimensional IRT, as its primary purpose is to test one or more models suggested by the multidimensional IRT analysis and it does not allow for item removal (Hair et al., 2019).

Confirmatory Factor Analysis (CFA). CFA was run only on the testing dataset after the misfitting items identified by the multidimensional IRT process were removed. CFA had the purpose of supplementing the evidence of structural validity generated by multidimensional IRT. The best models identified based upon results from multidimensional IRT were also tested and compared in CFA. This study ran an ordinal CFA with robust diagonally weighted least squares to compare the competing models (Desjardins & Bulut, 2018) in IBM Amos (Arbunckle, 2019). The cutoff values suggested by Hu and Bentler (1999) and Schreiber et al. (2006) were used to evaluate the models. The suggested cutoff values are chi-square/df < 3, CFI > .95, TLI > .95, RMSEA < .08, and SRMR < .05.

Reliability. To ensure the internal consistency of each factor and the accuracy of items, three different reliability values were computed. Cronbach's alpha was used to generate reliability values for each factor. Reliability values from IRT, namely person/estimated *a posteriori* plausible value reliability (EAP/PV) and item separation reliabilities, were also computed. According to Boone et al. (2014), person reliability refers to whether the "test

discriminate[s] the sample" based on their abilities, and item reliability refers to whether the sample is "big enough to precisely locate the items on the latent variable [construct]" (p. 230). Linacre (2012) indicated that the same cutoff values as Cronbach's alpha can be used to assess person and item reliabilities. Therefore, the >.70 cutoff suggested by DeVellis (2017) was used in this study for all three reliability tests.

Predictors

Demographics

In the demographic section of the instrument, students were asked about their gender identity, age, prior exposure to computer programming, and their confidence with using a computer (using a Likert scale ranging from 1-not confident at all to 10-strongly confident). They were also asked about their parents' terminal education, status of family income, and whether they had a personal computer at home. Cluster analysis was used to determine students' SES levels based on these later three variables.

Science and CS Attitudes

The science attitudes survey was adapted from an instrument developed by Unfried et al. (2015). The instrument originally consisted of nine items, but two items were misfitting based on the MNSQ values. Therefore, we proceeded with seven items for further analyses. The CS attitudes instrument consisted of 9 items. We used the MG-CS Attitudes developed by Rachmatullah et al. (2020). All of the items in these instruments were Likert-type ranging from 1 (strongly disagree) to 5 (strongly agree). All the items went through back-translation (Brislin, 1970; Cha et al., 2007) and re-validation processes as they were originally developed in English and for American students. The translated versions had EAP/PV reliability and Cronbach's alpha

of .879 and .896 for Science Attitudes instruments, and .863 and .874 for CS Attitudes instruments.

Science, Mathematics, and CT Scores

For science and mathematics achievements, students were asked to voluntarily provide their end-semester grades from the most recently completed semester. For CS achievement, we used their CT skills measured by CTA-M (Wiebe et al., 2019). This assessment consisted of 25 multiple-choice questions that compose of six items from Bebras challenges and 19 items from Roman-Gonzalez et al.'s (2017) CTt that use a block-based computer programming language as the context of questions. Wiebe et al. validated this instrument using IRT methods. Similar to the approach we used for CS and Science Attitudes instruments, CTA-M had undergone back-translation and re-validation processes through Rasch modeling. The re-validation processes resulted in a total 23 usable items. CTA-M had EAP/PV reliability and Cronbach's alpha values of .654 and .678, respectively.

Data Collection Procedure

The data collection process was done on three different days for students who were not involved in computational modeling activities and over five days for those who were involved in the activities. This separation of days was intended to reduce students' cognitive fatigue, which might impact their answers to the survey (Groves et al., 2011; Tourangeau et al., 2000). All the instruments were administered through a web-based survey using Qualtrics. For students who were not involved in the computational modeling activities, on the first day they took the demographics and science and CS attitudes survey, the second day they took the computational thinking assessment, and on the third day they took CISCI Instrument. For students involved in the activities, they took the same surveys before the activity with additional two days after the

activities where they took computational thinking assessment and attitudes instruments on the first day and CISCI instrument on the second day after the activities.

Data Analysis

Before performing further analyses, students' scores on science attitudes, CS attitudes, CT skills, and the five constructs in the CISCI instrument were converted from ordinal-type data to a ratio-interval (logit) scale through Rasch analysis using ConQuest 5.12.3. We used linear regression analysis to answer the second research question and to address the criterion validity suggested by AERA et al. (2014). The five constructs in the CISCI Instrument were treated as the dependent variables, while students' demographics information, and other attitudinal and cognitive measures were used as the predictors. Before running the regression test, we performed Pearson and Spearman correlation tests to explore the magnitude association between constructs in CISCI and the predictors. Pearson correlation tests were run on the CISCI constructs and the seven predictors (age, science scores, mathematics scores, CT scores, confidence using computer, and science and CS attitudes). Spearman correlation tests were performed to examine the correlation between CISCI constructs and gender, prior CS experience, and SES Levels. Moreover, we used paired-sample *t*-test to examine the impact of computational modeling activity of students' scores on the five CISCI constructs.

Findings

RQ 1. The Validation of CISCI Instrument

Content Validity

A total of eleven expert panels reviewed and assessed the original 17 items. Three items, ASE_03, CIN_01, and PSR_03, had CVR values below the cutoff. We then revised and kept the items in the instrument. The final revised version of the three items, the remaining items, and

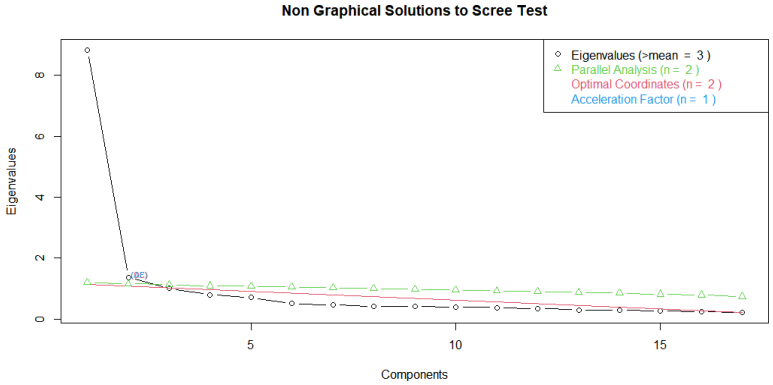
their corresponding CVR before revision are presented in Table 1. The initial content validity index (CVI) was .820, which is above the cutoff (Ayre & Scally, 2014; Lawshe, 1975).

Construct Validity

EFA. In the preliminary analysis addressing the assumptions for EFA, we found that the dataset we used for running EFA had a KMO value of .950. We also found a significant result from Bartlett's test of sphericity (X^2 = 6384.66, df = 136, p <.001). These results indicated that we met all the assumptions for conducting EFA. We then proceeded with conducting EFA with two different approaches, using parallel analysis and eigenvalue cutoff of < 1. The results are shown in the scree plot in Figure 3.

Figure 3

Scree Plot



The scree plot shows that parallel analysis and eigenvalues < 1 generated two different factor solutions. Based on the parallel analysis, a two-factor solution was identified in our data. The first factor consisted of only CIS Career Value items, and the rest of the items were loaded into the second factor. The total variance explained was 55%, with the first and second factors accounted for 37% and 18%, respectively. The results based on eigenvalues < 1 suggested a

three-factor model consisting of (1) CIS Career Value, (2) CIS Parental Supports, and Role Models, and (3) the third factor is composed of the remaining three constructs. This approach's solution had a cumulative variance of 59% with 15%, 21%, and 22% accounted by CIS Career Value, CIS Parental Supports and Role Models, and the third factor, respectively. All the factor loadings from these analyses are presented in Table 1.

Multidimensional IRT. We then compared four different models using multidimensional IRT analysis to gather more evidence for how each item behaves within each model. These four comparable models were one-dimensional, two-dimensional, and three-dimensional models based on EFA results and a five-dimensional model based on our initial conceptualization. Table 2 shows the results.

We found that PSR_02 and PSR_03 were identified as misfitting items with unweighted and weighted MNSQ values above 1.40 in the four comparable models. Of note, PSR_03 was also identified as having a CVR value below the cutoff in its initial wording. These two items were then removed from the models, and multidimensional IRT analysis was rerun. As shown in Table 2, the one-dimensional and two-dimensional models still contained misfitting items, while the three-dimensional and five-dimensional models were free from any misfitting items. Moreover, the three- and five-dimensional models also had lower final deviance (FD), AIC, and BIC values than one- and two-dimensional models. These lower values and no misfitting items indicated that the three- and five-dimensional model fit our data better than the other two models. Therefore, we proceeded with these three and five-dimensional models with 15 items (referred to as CISCI 2.0) for further analyses.

Table 1

CISCI 1.0 and its CVR and EFA Factor Loadings

Construct	Code	Item (Revised/Final)	CVR (≥ .818; Before Revision)	Parallel Analysis			Eigenvalue < 1			
				F1	F2	U	F1	F2	F3	U
CIS Career Value	CV_01	I believe jobs like astronaut, doctor, scientist, roboticist, and computer engineer are important because they combine science and computer science (coding)	0.818		.742	.400			.727	.395
	CV_02	I believe jobs that combine science and computer science (coding) help me come up with new ideas.	1.000		.834	.230			.810	.230
	CV_03	I believe jobs that combine science and computer science (coding) provide opportunities to collaborate with others.	0.818		.725	.345			.695	.345
CIS Academic Self-efficacy	ASE_01	I am confident that I will do well in classroom activities that combine science and computer science (coding).	0.818	.597		.513	.626			.454
	ASE_02	I am confident that I can learn skills that combine science and computer science (coding).	1.000	.615		.485	.680			.393
	ASE_03	I am confident that I can understand science through computer science methods (coding)	***0.636***	.709		.384	.727			.301
CIS Career Self-efficacy	CAS_01	I am confident I can work in a field that combines science and computer science (coding).	1.000	.778		.323	.563	.552		.332
	CAS_02	I am confident that I can be successful in a field that combines science and computer science (coding)	0.818	.776		.301	.595	.520		.311
	CAS_03	I am confident I can combine science and computer science (coding) in my future career.	0.818	.730		.388	.543	.503		.398
CIS Career Interest and Goal	CIN_01	I believe that combining scientific concepts and computer science would be useful after I am done with schooling.	***0.455***	.628		.465	.604			.435
	CIN_02	I want to enter a career that combines science and computer science (coding).	1.000	.700		.413	.421	.596		.391
	CIN_03	I would feel comfortable talking to people who combine science and computer science (coding) in their work.	0.818	.717		.388	.570	.472		.387
CIS Parental Supports and Role Models	PSR_01	I would talk to my parents about my goal to have a job that combines science and computer science (coding).	0.818	.601		.521		.623		.435
	PSR_02	I believe that my parents would support me in the future if I choose a career that combines science and computer science (coding)	0.818	.523		.573		.459		.554
	PSR_03	I know people in my family who combine science and computer science (coding) in their work.	***0.455***	.570		.640		.567		.580
	PSR_04	Someone I know who has a career that combines science and computer science (coding) inspires me to have their job.	0.818	.578		.644		.641		.530
	PSR_05	I have role models who use coding in their science careers.	1.000	.576		.650		.614		.560

Notes: F = Factor; U = Uniqueness; the items presented here went through back-translation process where a native Indonesian speaker who is fluent in English translated the items from Indonesian to English, and then two English native speakers checked the translation for accuracy

Table 2

Comparing the four models of CISCI Instrument

Model	17 Items						15 Items					
	X^2	*df*	FD	AIC	BIC	Total Misfitting Items	X^2	*df*	FD	AIC	BIC	Total Misfitting Items
One-dimensional	2677.50	16	42077.48	42119.48	42228.18	2	2514.57	14	36048.02	36086.02	36184.37	2
Two-dimensional	858.23	15	41336.36	41382.36	41501.41	2	678.94	13	35173.27	35215.27	35323.98	2
Three-dimensional	1285.57	14	40867.89	40919.89	41054.48	2	789.63	12	34801.28	34849.28	34973.51	0
Five-dimensional	1255.69	12	40694.86	40764.86	40946.03	2	703.99	10	34496.21	34562.21	34733.03	0

We then further examined the quality of items within these three- and five-dimensional models. Table 3 presents the MNSQ values for each item in these two models. As expected, both the unweighted and weighted MNSQ values for all the items in these two models are within the range of acceptable values 0.6–1.4 (Wright & Linacre, 1994). This indicated that the items could differentiate students with high interest in computationally intensive science careers from those who had low interest in such careers. The Wright Maps for these two models showing the range of students' interest and difficulty levels for each item are depicted in Figure 4. Next, we performed CFA to examine the structure of these two models and to compare whether there is a significant difference in the quality of these two models.

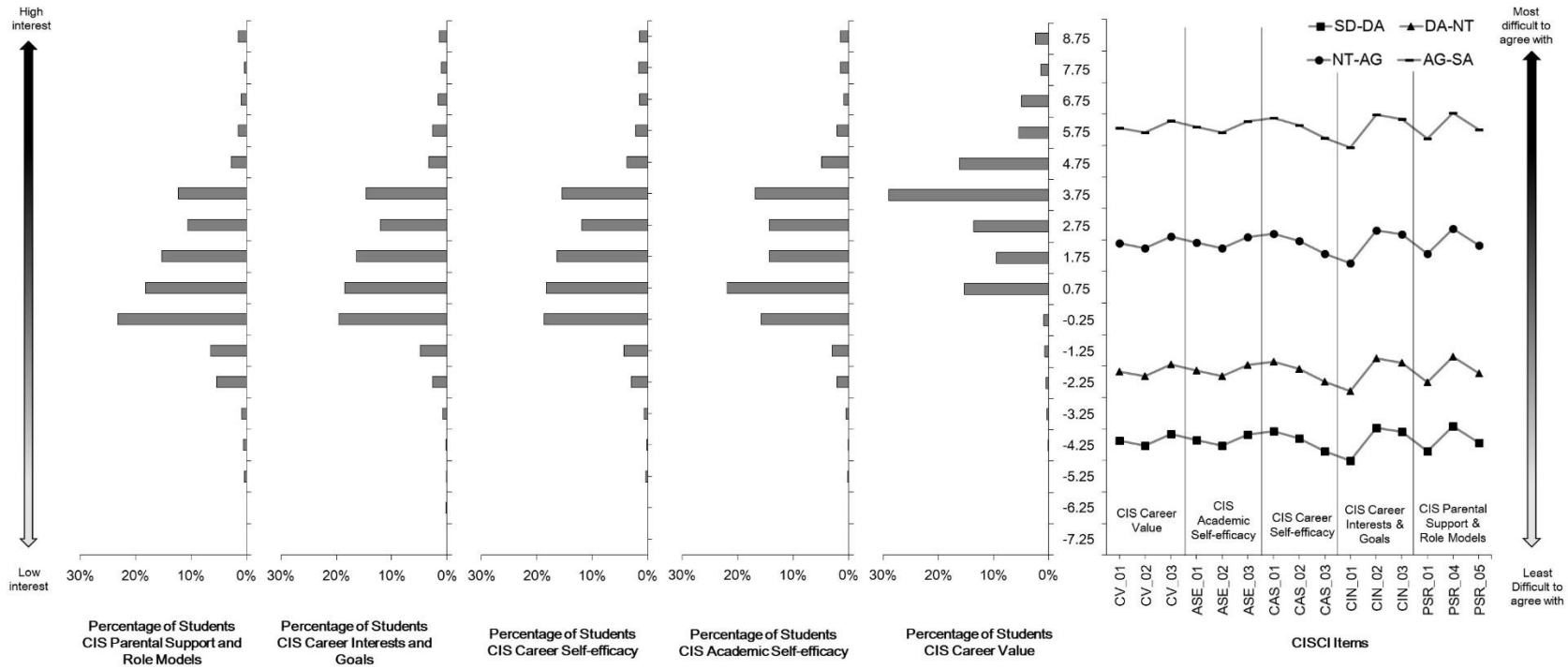
Table 3

Comparing Item Quality and Difficulty between CISCI 2.0 Three- and Five-Dimensional Models

Dimension	Item Code	3D 15 Items				5D 15 Items			
		Estimate	Un-weighted MNSQ	Weight-ed MNSQ	α if item deleted	Estimate	Un-weight-ed MNSQ	Weight-ed MNSQ	α if item deleted
Career value	CV_01	-0.023	1.02	1.04	.818	-0.024	1.00	1.04	.818
	CV_02	-0.161	0.82	0.84	.740	-0.167	0.78	0.82	.740
	CV_03	0.185	0.83	0.86	.790	0.192	0.82	0.84	.790
Academic self-efficacy	ASE_01	-0.113	1.12	1.15	.918	-0.002	1.05	1.06	.814
	ASE_02	-0.283	0.98	0.99	.917	-0.177	0.86	0.87	.761
	ASE_03	0.062	0.91	0.94	.914	0.180	0.86	0.87	.774
Career self-efficacy	CAS_01	0.317	0.82	0.84	.913	0.289	0.83	0.85	.816
	CAS_02	0.096	0.83	0.85	.911	0.061	0.83	0.84	.806
	CAS_03	-0.300	0.90	0.93	.915	-0.350	0.91	0.93	.838
Career interest and goal	CIN_01	-0.549	1.13	1.16	.918	-0.646	1.19	1.19	.747
	CIN_02	0.449	1.04	1.05	.917	0.389	1.04	1.05	.748
	CIN_03	0.322	1.00	1.03	.916	0.257	1.01	1.03	.667
Parental support and role models	PSR_01	-0.348	1.20	1.21	.716	-0.360	1.27	1.29	.716
	PSR_04	0.432	1.22	1.19	.656	0.446	1.28	1.25	.656
	PSR_05	-0.084	1.08	1.08	.662	-0.085	1.14	1.14	.662

Figure 4

The Wright Map of the CISCI 2.0



Notes: The agreement difficulties for each item on the right are based upon Thurstonian thresholds. The Thurstonian thresholds specify a location where a student who has the same logit score as the threshold has a 50% probability of selecting a given Likert scale. Students' career interests are depicted with histograms on the left. When a student is precisely at the Thurstonian threshold, the student has an equal probability of choosing the Likert scale above or below the threshold. SD= Strongly Disagree, DA=Disagree, NT=Neutral, AG=Agree, SA=Strongly Agree.

CFA. We ran CFA on the testing dataset to compare the statistical structure of the three- and five-dimensional models (we used factor instead of dimension in CFA). Table 4 shows the results comparing the indices between the three- and five-factor models. Based on the results, we found that the five-factor model had higher CFI and TLI values and lower RMSEA and SRMR than the three-factor model. Moreover, the five-factor model had indices that were above the cutoff values. This indicated that the five-factor model was a better model for our data. The structure of this five-factor model is depicted in Figure 5. Furthermore, we used the five-factor model to compute the reliability values.

Table 4

Results from CFA Comparing the Three- And Five-Factor Models

Indicator (cutoff)	3-factor Model	5-factor Model
X^2	569.60	313.78
df	87	80
X^2/df	6.55	3.92
p-value	< .001	< .001
CFI (> .95)	.92	.96
TLI (> .95)	.90	.95
RMSEA (< .08)	.09	.06
RMSEA 90CI	[.082, .096]	[.057, .072]
SRMR	.048	.038
ΔX^2 (Δdf)	-	255.82 (7)
p-value for ΔX^2	-	< .001

Reliability. The five-factor model consisted of five constructs: (a) CIS Career Value, (b) CIS Academic Self-efficacy, (c) CIS Career Self-efficacy, (d) CIS Career Interest and Goal, and (e) CIS Parental Supports and Role Models. These constructs had a separation reliability value of .991. The five constructs had EAP/PV values of .824, .904, .915, .916, and .912, respectively. The Cronbach's α for the five constructs were .844, .844, .872, .796, and .760. All of these

reliability values are within the range of acceptable values (.70; DeVellis, 2018). Therefore, each of the constructs is internally consistent.

Figure 5

Structural Visualization of the Five-Factor Model of CISCI 2.0

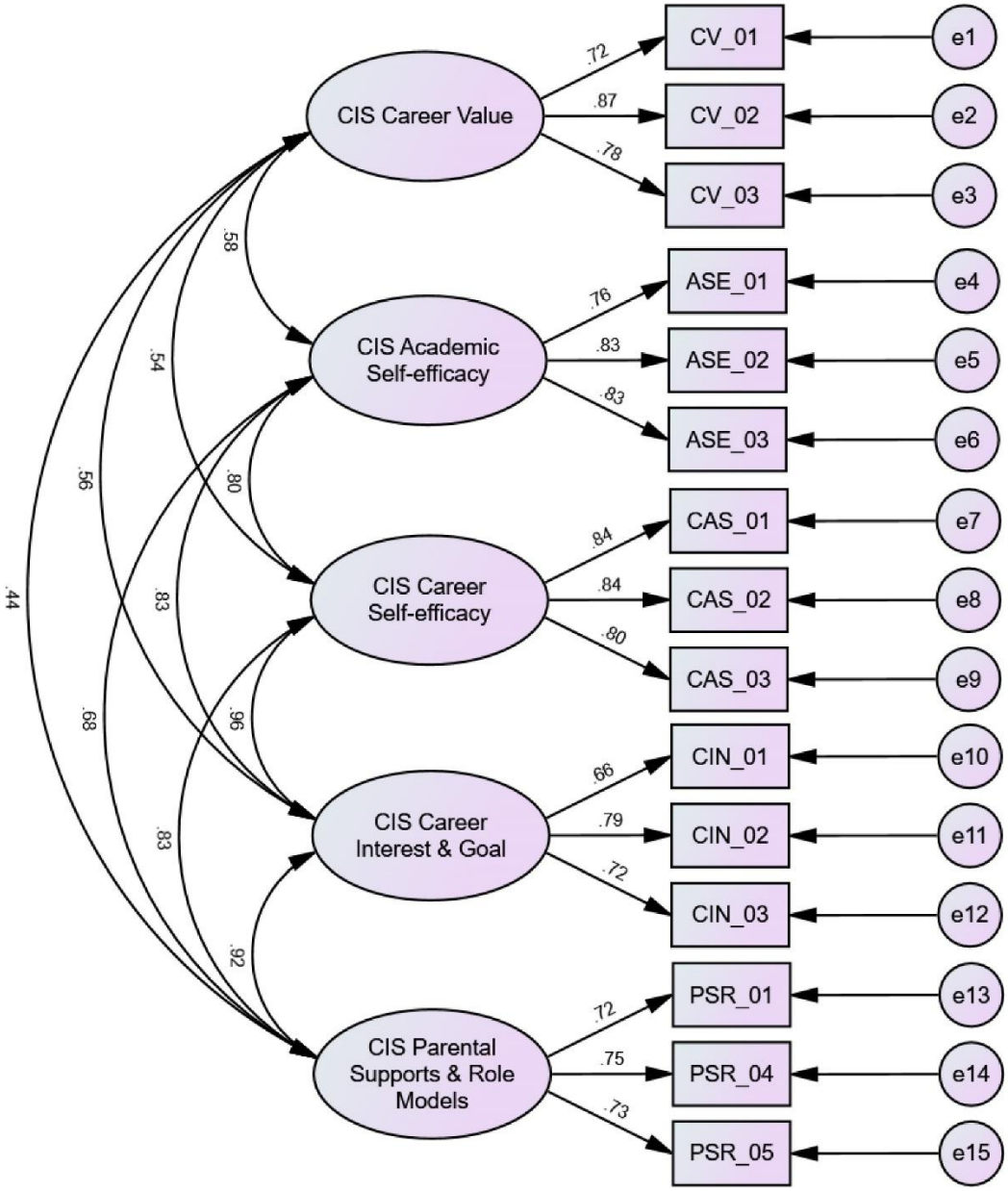


Table 5

Pearson Coefficient Correlations

Variable		Variable														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CIS Career Value	1	-	.804**	.759**	.762**	.644**	.101**	.132**	.236**	.578**	.620**	.236**	.058	.130**	.037	.024
CIS Academic Self-efficacy	2			.956**	.958**	.854**	-.018	.004	.084*	.716**	.722**	.347**	.001	.152**	-.112**	.010
CIS Career Self-efficacy	3				.997**	.941**	-.035	-.001	.065	.733**	.726**	.310**	-.017	.151**	-.123**	.034
CIS Career Interest and Goal	4					.936**	-.035	.000	.066	.735**	.724**	.310**	-.01	.158**	-.121**	.029
CIS Parental Supports	5						-.096**	-.062	.007	.667**	.661**	.275**	-.045	.137**	-.142**	.044
Science Scores	6							.724**	.389**	.017	.004	-.022	.253**	.034	.459**	.032
Mathematics Scores	7								.372**	.042	.005	.015	.245**	.027	.485**	.077*
CT Scores	8									.093*	.111**	.070	.224**	.130**	.323**	-.036
Science Attitudes	9										.611**	.302**	-.01	.113**	-.048	.108**
CS Attitudes	10											.390**	.02	.151**	-.106**	-.024
Confidence with Using Computer	11												-.041	.228**	.154**	-.086*
Age	12													.069*	.051	-.032
Prior CS Experience[S]	13														.154**	-.086*
SES[S]	14															-.047
Gender (Female coded as 2) [S]	15															–
M (Median for Variables 13, 14)		2.950	1.482	1.319	1.270	0.740	84.680	82.912	-0.421	1.413	1.397	2.978	12.849	2	3	-
SD		1.801	1.980	2.142	2.095	2.030	8.787	11.438	0.496	1.882	1.270	0.747	0.878	0.936	0.765	-
Skewness		0.237	0.600	0.556	0.579	0.555	-1.539	-1.711	0.515	0.248	0.098	-0.016	0.774	0.676	-0.787	-
Kurtosis		0.678	0.897	0.713	0.774	1.193	8.345	6.524	0.088	0.200	0.525	0.581	0.353	-0.320	-0.863	-

Note: [S] = Spearman rho's coefficient correlation; no asterisk $p > .05$, * $p < .05$, ** $p < .01$; Variables 1-5 and 8-10 are in the form of logit scores

RQ2. Predictors of Students' CISCI

We then ran multiple linear regressions to answer the second research question about the significant predictors of students' interests in computationally intensive science careers. These predictors were determined based on SCCT interest model and previous studies suggesting that students' career interests are predicted by their cognitive, affective, and demographic factors. These regression analyses were also intended to address the criterion validity of the CISCI instrument. Prior to running the regression test, correlation tests were run. The results are shown in Table 5. We found that no predictor was highly correlated (> .80) with the five constructs in the CISCI instrument, meaning no multicollinearity was found. Thus, we could proceed to perform multiple linear regressions for each of the five constructs. The results are displayed in Table 6.

CIS Career Value

The results showed that the model explained 52% of the variance in CIS Career Value, $R^2_{adj} = .52$, $F(10, 664) = 74.12$, $p < .001$. We found that students' CS and Science attitudes were significantly and positively associated with the CIS Career Value. Moreover, students' mathematics and CT scores were also found to significantly and positively predict students' CIS Career Value. Regarding students' demographics, results revealed that gender, SES, and prior CS experience were not significant predictors of students' CIS Career Value. Interestingly, the science score and confidence with using computers were also not significant predictors of students' CIS Career Value.

CIS Academic Self-efficacy

The model explained 67% of the variance in CIS Academic Self-efficacy, $R^2_{adj} = .67$, $F(10, 664) = 135.43$, $p < .001$. Similar to CIS Career Value, students' CS and Science attitudes

were positively associated with students' CIS Academic Self-efficacy. Other affective, cognitive, and demographic variables did not significantly predict students' CIS Academic Self-efficacy, except students' SES. We found that SES was negatively associated with students' CIS Academic Self-efficacy. This means that students with a higher SES had significantly lower CIS Academic Self-efficacy than students with lower SES levels.

Table 6

Results from Linear Regression for the Five Constructs In CISCI with Standardized Beta Values

Predictor	CIS Career Value		CIS Academic Self-efficacy		CIS Career Self-efficacy		CIS Career Interest and Goal		CIS Parental Supports and Role Models	
	B	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*
Age	-0.018	.521	-0.055	.321	-0.023	.314	-0.021	.372	-0.025	.350
Gender	-0.008	.781	-0.137	.128	-0.022	.324	-0.024	.284	-0.017	.514
SES	0.041	.177	-0.158	.015	-0.053	.033	-0.055	.028	-0.059	.042
Prior CS Experience	0.018	.518	0.096	.054	0.057	.016	0.061	.010	0.082	.003
Science Attitudes	0.332	< .001	0.506	< .001	0.496	< .001	0.499	< .001	0.451	< .001
CS Attitudes	0.438	< .001	0.636	< .001	0.402	< .001	0.397	< .001	0.365	< .001
Confidence with Using Computer	-0.044	.147	0.077	.239	-0.023	.358	-0.022	.376	-0.029	.314
Science Score	-0.019	.617	-0.003	.704	-0.030	.354	-0.035	.273	-0.047	.205
Mathematics Score	0.091	.021	0.009	.115	0.052	.109	0.056	.082	0.017	.650
CT Score	0.125	< .001	0.038	.705	-0.009	.729	-0.007	.786	-0.041	.160

CIS Career Self-efficacy

The regression model explained 67% of variance in CIS Career Self-efficacy, $R^2_{adj} = .67$, $F(10, 664) = 140.39$, $p < .001$. The same significant predictors for CIS Academic Self-efficacy were found in the regression model of CIS Career Self-efficacy. Science and CS attitudes were found to be positively associated with students' CIS Career Self-efficacy. We also found that SES was negatively associated with students' CIS Career Self-efficacy. In addition, the results revealed that Prior CS Experience was a significant positive predictor of students' CIS Career

Self-efficacy. The other variables were found to not be significant predictors of students' CIS Career Self-efficacy.

CIS Career Interests and Goals

The model explained 68% of variance in CIS Interest and Goal, R^2_{adj} =.68, $F(10, 664)$ = 140.92, $p < .001$. Science and CS attitudes were also found to be significantly associated with CIS Career Interests and Goal. Like in the two previous constructs, Prior CS Experience was also a significant positive predictor of CIS Career Interests and Goal. Moreover, SES was found to be negatively associated with students' Career Interests and Goal.

CIS Parental Supports and Role Models

Last, the regression model explained 56% of variance in CIS Parental Supports and Role Models, R^2_{adj} = .56, $F(10, 664) = 87.32$, $p < .001$. The results showed that CS and Science attitudes were significant positive predictors of this construct, so were Prior CS Experience. We found that SES level was a significant negative predictor of students' CIS Parental Supports and Role Models. These results suggest that students with lower SES statuses reported discussing CIS careers more frequently than those with higher SES statuses.

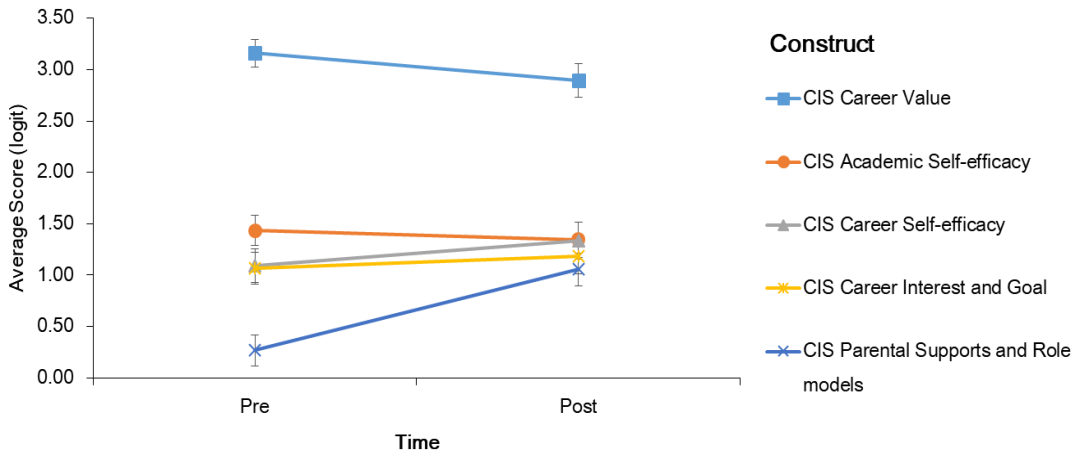
RQ3. The Impacts of the Scientific Computational Modeling Activity on Students' Career Interests

A paired-sample *t*-test was used to examine the changes in students' scores on the five constructs before and after the scientific computational modeling activity. We found no significant changes in students' CIS Career Value: $t(185) = 1.81$, $p = .071$, $d = 0.13$; CIS Academic Self-efficacy: $t(185) = 0.64$, $p = .524$, $d = 0.04$; CIS Career Self-efficacy: $t(185) = -1.54$, $p = .125$, $d = 0.10$; and CIS Career Interest Goal: $t(185) = -0.79$, $p = .434$, $d = 0.05$. However, we found a significant increase in students' scores on CIS Parental Supports from

before the activity to after the activity: $t(185) = -5.18$, $p < .001$). This significant change was with small effect size ($d = 0.37$). The overall results are visualized in Figure 6.

Figure 6

Changes in Students' Scores on CISCI Instrument before and after A Scientific Computational Modeling Activity



Discussion and Implications

The current study's findings document that the CISCI instrument was multidimensional, consisting of five constructs: career value, academic self-efficacy, career self-efficacy, career interests and goals, and parental support and role models. The results also revealed that the following five constructs shared common predictors: science and CS attitudes, prior CS experience, and socioeconomic level. Furthermore, the current study also found that participating in a computationally rich science activity for the first time might increase middle-school students' frequency of discussing CIS careers with their parents.

The CISCI Instrument Validation

The results from Multidimensional IRT analysis and CFA showed that the CISCI instrument comprised five constructs, whereas EFA indicated two- and three-factor solutions. These discrepancies in the number of factors might be due to the relatively high correlation

among constructs reported in Table 5: specifically, CIS academic self-efficacy, career self-efficacy, career interests and goals, and parental support and role models. Such strong correlations among constructs might have prevented EFA from detecting more factors, given EFA is based on correlations among items (Hair et al., 2019). This issue was also found in Smith's (1996) study on identifying the best dimensionality approach by comparing IRT analysis to factor analysis. Smith found that when the instrument consisted of uncorrelated factors, EFA worked best in identifying the numbers of items. However, when highly correlated items and factors dominated the instrument, Smith suggested using IRT analysis. Smith's findings resonate with our findings presented in Table 5. The four factors were highly correlated, which might have resulted in EFA generating two- and three-factor solutions. Such highly correlated items or factors might be attributed to middle-school students' inability to differentiate the latent constructs: academic self-efficacy, career self-efficacy, and career interests and goals. This pattern was also found in our previous study on validating a CS-attitude instrument for middle-school students in the US, where students' responses were unable to differentiate CS self-efficacy and outcome expectations based on EFA (Rachmatullah et al., 2020). However, IRT analysis and CFA determined the numbers of latent factors more efficiently. Christensen et al. (2012) also suggested the combination of IRT analysis and CFA, especially for studying dimensionality with smaller numbers of items. Therefore, again, this study's results showed the superiority of the combination of IRT analysis and CFA in determining the instrument's dimensionality, item difficulty, and item quality with a relatively smaller number of items.

The construct separation findings align with previous studies that utilized SCCT to develop and validate an instrument to measure students' STEM career interests, especially Roller et al. (2020) and Shin et al. (2018). Although these two studies only used either EFA and CFA or

multidimensional IRT, both found the same construct separation as the current study. Shin et al.'s (2018) study also revealed a relatively high correlation among the constructs in their instrument, but not as strong as in the current study (< .80). It is noteworthy that Shin et al. used data from the combination of middle and high school students' populations; thus, developmental-related factors might have influenced the results. Therefore, future studies are encouraged to gather more evidence for the CISCI construct separation from older-aged students. Moreover, Kier et al. (2014) had the same student populations as the current study and found a one factor solution. However, they only utilized CFA to gather the construct validity evidence for their instrument. Hence, we believe that our study has corroborated these previous findings by performing more thorough and robust analyses to collect the construct validity evidence.

The CISCI instrument has practical implications, as it has a stable and smaller number of items. The number of items measuring each construct was intentional, given that we wanted a lean instrument to measure students' CIS career interest in a relatively short time. This approach can also prevent students' testing fatigue if researchers or policymakers administer other instruments together with the CISCI instrument. Moreover, CISCI's constructs can be used independently, given that results from IRT analysis and CFA revealed five different constructs. Thus, researchers and policymakers can choose whichever construct applies to their study. Furthermore, future studies can still use the three-dimensional model of CISCI given that our multidimensional IRT results showed this model was also stable. In addition, the regression results revealed that some of the constructs shared common predictors.

CISCI Predictors and Intervention

Multilinear regression analysis results revealed that the five CISCI constructs shared four predominantly common significant predictors: science attitudes, CS attitudes, SES, and prior CS

experience. As predicted, science and CS attitudes are highly associated with and significantly predictive of all CISCI instrument constructs. This finding aligns with Lent et al.'s (2002) SCCT of interest model that shows attitudes in the forms of self-efficacy and outcome expectations are the root of career interests. This finding implies that both science and CS attitudes are equally integral in forming students' interests in computationally rich science careers. As shown in Table 6, beta values from regression analysis for these two attitudinal constructs were relatively the same or close.

The dominant power of science and CS attitudes in predicting students' interests might have made the cognitive factors—science, mathematics, and CT scores—insignificant in predicting students' interests in CIS careers. This finding was unexpected, as previous studies have shown that cognitive achievements are significant predictors of career interests (Dabney et al., 2012; Hazari et al., 2013). Such results may occur when attitudinal factors are not measured or included in the model. As Lent et al. (2002) described in the SCCT interest model, cognitive achievements, which may be part of learning experiences, do not directly relate to career interests. The relationships are mediated by self-efficacy and outcome expectations, which are the constructs within the science and CS-attitude instruments used in the current study. Therefore, the present study results again show the alignment with Lent et al.'s SCCT, indicating the possibility of science and CS attitudes mediating cognitive performances and CIS career interests. This observation does not indicate that cognitive performances are not crucial in influencing students' career interests, but that academic performance may strengthen students' attitudes toward science and CS. This relationship can be confirmed by performing a path analysis as shown in Smith's (2002) study on the relationships between information technologies related to cognitive performance, career interests, and self-efficacy and outcome expectations.

Smith found that self-efficacy and outcome expectancy significantly mediated the relationship between students' cognitive performance and career interests. Hence, our findings support the idea that involving students in activities that nurture positive science and CS attitudes is a strong strategy for improving their interest in computationally intensive science careers. Such activities may include scientific computational modeling activities where students model scientific phenomena using computer programming.

Engaging students in a computationally rich environment where they are involved in using CS-related methods and tools may also explain why prior CS experience is a significant factor of students' interests in CIS careers. This finding may connect to Bandura's (1997) self-efficacy theory about past or mastery experience as a source of students' self-efficacy, which is part of students' attitudes. Prior experiences are significant predictors of students' attitudes toward science and CS (Hinckle et al., 2020; Jones et al., 2021; Leonard et al., 2016). Our finding has extended the understanding of the prior experience–attitudes relationship by showing that prior experience is a significant predictor of career interests. Jones et al. (2021) provided reasoning to this prior experience–career interest relationship, offering that prior experience may have exposed students to tools and people working in such careers, which eventually helped students to internalize these experiences as their interests. This reasoning resonated with our intervention results, where we found a significant increase in students' CIS parental support and inclusion of role models. Engaging students in computationally rich science activities might have increased the frequency of student–parent discussions surrounding CS-integrated science subjects and careers. These discussions might include students asking whether they have someone in their immediate or non-immediate family who works in such sectors. In addition, there is a possibility that the significant increase in such parent-student discussions was due to

the online learning environment. Thus, the parents had more opportunities to observe and talk to their children. However, we lacked such data, given that we did not collect any interview data from either students or parents. Therefore, we recommend future studies to gather such data to substantiate this finding and may frame the study using the family habitus theory (Archer et al., 2012).

Family's characteristics may also explain the significant negative impacts of students' SES level on their CIS career interests. This finding aligns with discussions from previous studies contending that students from lower-income families may perceive that CIS or other STEM careers offer viable options for increasing their social and/or economic status (Melguizo & Wolniak, 2012). As raised in previous studies (Jones et al., 2017; Settlage & Meadows, 2002)—regardless of access discrepancies to tools or science resources, CIS or STEM-related careers between lower- and higher-SES students—students from lower-SES families (especially in Indonesia) might have considered CIS careers as well-respected and high-paying jobs that could elevate their family's SES status. Hence, students from lower-SES families tended to have higher interest and aspirations regarding such careers. More in-depth studies are needed to delve into this reasoning, including future studies utilizing qualitative data on what students and parents, especially those from lower-SES groups, think about CIS-related careers.

We identified an interesting finding regarding predictors of CIS career values that differ from the other four constructs: Mathematics and CT scores were significant predictors of students' CIS career value, although the two were not significant predictors of the other four CISCI constructs. This finding implies that the higher students' scores on mathematics and CT, the more they recognize CIS careers' importance. Possible reasoning for this relationship is that both CS and CT share some common principles with, or are built on, mathematical theories and

thinking (Papadimitriou, 2003; Salac et al., 2020). Students with higher mathematics achievement perceive mathematics as an essential subject that has relevancy to many problem-solving skills (Flegg et al., 2012; Lucas & Fugitt, 2007). Such perceptions of relevance and importance of mathematics might have influenced students in the current study to consider CIS careers as more valuable, making CT and mathematics scores the significant predictors.

Last, the intervention results showed that the impact of scientific computational modeling activities was significant only on students' CIS parental support and the inclusion of role models, as mentioned above. However, we did not find a substantial impact on the four other CISCI constructs. This finding was unexpected because we hypothesized a possible increase in all the CISCI constructs after the activities. We believe that we did not see a significant increase in other constructs because, although some students had prior CS experiences, they did not have rich experience in learning science using CS methods and tools, such as computer programming. Such first encounters might have masked the impact of the intervention, as students were largely engaged in more low-level tasks of learning these new tools and approaches. Schools and teachers should advocate more extended scientific computational modeling activities via after-school and extracurricular activities or integrate CT-CS into other subjects to render longer-term impacts of such activities on students' CISCI and their learning. Such extended activities can also help researchers and policymakers trace and evaluate students' CISCI development to more fully understand the activities and students' career interests in CIS.

Conclusion, Limitations, and Future Directions

This study developed and validated an instrument, CISCI, to measure middle-school students' interests in computationally intensive science careers. This study documented that students' interests in such jobs are predicted by their science and CS attitudes, prior CS

experience, SES level, and CT and mathematics scores. We also found a significant impact of computationally rich science activities on students' parental support and role models. Our findings extend the literature on students' career interests in computationally intensive science, a topic that has not yet been explored. Situating this study in Indonesia has both helped generalize findings largely coming out of Western countries, while also highlighting potentially unique relationships between SES and computationally intensive science careers.

Like many other studies, the current study has some limitations. First, the students participating in this study were not randomly selected; instead, they were included based upon the schools' willingness to participate. Even though the gender proportion matched Indonesia's overall demographics profiles, the SES proportion did not match the national profiles, which might have biased results. Second, the intervention focused primarily on a single grade, seventh graders, thus limiting the studies' generalization across the middle grades years. Furthermore, the intervention was conducted in an online synchronous learning environment due to the COVID-19 global pandemic; only students possessing computers could join. We acknowledge this condition as a study limitation, as it is likely skewed towards students from higher SES-level families.

Future studies are encouraged to address the current study's limitations, such as offering scientific computational modeling activities in a face-to-face setting and specifically recruiting from low-SES contexts. More extended activities are highly encouraged to explore further the trajectory of students' career interests in CIS. Additional studies can also compare various computational modeling activities in terms of the platform and programming language and different instructional strategies, such as comparing the UMC strategy to others. Regarding CISCI instrument validation, given that career interests are contextual, more evidence from

different grade levels (e.g., high school) and different countries is crucial to adding evidence towards the instrument's criterion validity. Future studies can also use a different instrument to measure CT, such as CT disposition (Hava & Ünlü, 2021), to confirm whether results would be the same as when CT is measured using cognitive assessments, as in this study.

Acknowledgement

We thank all the students and teachers participating in this study. We thank Jennifer Houchins, Cansu Tatar, and all the expert panels who provided significant feedback on the first version of the CISCI instrument. We also thank Dr. Jessica Vandenberg and Dolly Bounajim for providing thoughtful feedback on the initial draft of this work. This material is based upon the work supported by the National Science Foundation under Grant No. DRL-1640141. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The portions of this study were also supported by NC State University's College of Education and the Friday Institute for Educational Innovation through Dissertation Research Continuity Award and Graduate Student Fellow Award.

REFERENCES

Adams R. J., & Wu, M. (2010). *Multidimensional model. (August 2010)*. Retrieved on March 10, 2019 from https://www.acer.org/files/Conquest-Tutorial-7-MultidimensionalModels.pdf

Adams, R. J., Wu, M., Macaskill, G., Haldane, S., Sun, X. X., & Cloney, D. (2020). *ACER ConQuest: Generalised item response modelling software [Computer software]. Version 5.12.3*. Camberwell, Victoria: Australian Council for Educational Research.

Aksit, O., & Wiebe, E. N. (2020). Exploring force and motion concepts in middle grades using computational modeling: A classroom intervention study. *Journal of Science Education and Technology*, *29*(1), 65–82. https://doi.org/10.1007/s10956-019-09800-z

Alshahrani, A., Ross, I., & Wood, M. I. (2018, August). Using social cognitive career theory to understand why students choose to study computer science. In *Proceedings of the 2018 ACM Conference on International Computing Education Research* (pp. 205–214). https://doi.org/10.1145/3230977.3230994

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

Arbuckle, J. L. (2019). *Amos (Version 25.0) [Computer Program]*. Chicago: SPSS.

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2012). Science aspirations, capital, and family habitus: How families shape children's engagement and identification with science. *American Educational Research Journal*, *49*(5), 881–908. https://doi.org/10.3102/0002831211433290

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, *47*(1), 79–86. https://doi.org/10.1177/0748175613513808

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice–Hall.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: W. H. Freeman and Company.

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

Boone, W. J. (2016). Rasch analysis for instrument development: why, when, and how?. *CBE—Life Sciences Education*, *15*(4), 1–7.

Boone, W. J., Staver, J. R., & Yale, M. S. (2013). *Rasch analysis in the human sciences*. Springer, Dordrecht.

Bortz, W. W., Gautam, A., Tatar, D., & Lipscomb, K. (2020). Missing in measurement: Why identifying learning in integrated domains is so hard. *Journal of Science Education and Technology*, *29*(1), 121–136.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: issues and techniques. *Journal of Advanced Nursing*, *58*(4), 386–395. https://doi.org/10.1111/j.1365-2648.2007.04242.x

Christensen, K. B., Engelhard, G., & Salzberger Jr, T. (2012). Ask the experts: Rasch vs. factor analysis. *Rasch Measurement Transactions*, *26*(3), 1373–1378. https://www.rasch.org/rmt/rmt263a.htm

Dabney, K. P., Tai, R. H., Almarode, J. T., Miller-Friedmann, J. L., Sonnert, G., Sadler, P. M., & Hazari, Z. (2012). Out-of-school time science activities and their association with career interest in STEM. *International Journal of Science Education, Part B: Communication and Public Engagement*, *2*(1), 63–79. https://doi.org/10.1080/21548455.2011.629455

Desjardins, C. D., & Bulut, O. (2018). *Handbook of educational measurement and psychometrics using R*. Boca Raton, FL: CRC Press.

DeVellis, R. F. (2017). *Scale development: Theory and applications (Vol. 26)*. Thousand Oaks, CA: Sage Publications.

Dickes, A. C., Kamarainen, A., Metcalf, S. J., Gün-Yildiz, S., Brennan, K., Grotzer, T., & Dede, C. (2019). Scaffolding ecosystems science practice by blending immersive environments and computational modeling. *British Journal of Educational Technology*, *50*(5), 2181–2202. https://doi.org/10.1111/bjet.12806

European Center for the Development of Vocational Training. (2014). *Rising STEMs*. Retrieved from https://www.cedefop.europa.eu/en/publications-and-resources/statistics-and-indicators/ statistics-and-graphs/rising-stems.

Falloon, G., Hatzigianni, M., Bower, M., Forbes, A., & Stevenson, M. (2020). Understanding K-12 STEM education: a framework for developing STEM literacy. *Journal of Science Education and Technology*, *29*(3), 369–385. https://doi.org/10.1007/s10956-020-09823-x

Flegg, J., Mallet, D., & Lupton, M. (2012). Students' perceptions of the relevance of mathematics in engineering. *International Journal of Mathematical Education in Science and Technology*, *43*(6), 717–732. https://doi.org/10.1080/0020739X.2011.644333

Funke, A., Berges, M., & Hubwieser, P. (2016, March). Different perceptions of computer science. In *2016 International Conference on Learning and Teaching in Computing and Engineering (LaTICE)* (pp. 14–18). IEEE. https://doi.org/10.1109/LaTiCE.2016.1

Goode, J., & Chapman, G. (2016). *Exploring computer science*. University of Oregon, Eugene, OR. Retrieved September 11, 2020 from https://www.teach21.us/uploads/1/3/0/5/13053428/samplecscurriculum.pdf

Grover, S., & Pea, R. (2018). Computational thinking: A competency whose time has come. In S. Sentence, E. Barendsen., & C. Schulte (Eds.) *Computer Science Education: Perspectives on Teaching and Learning in School* (pp. 19-38). London, UK: Bloomsbury Academic.

Groves, R.M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology: Second edition*. Hoboken, NJ: John Wiley & Sons, Inc.

Hair, J.F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis: Eighth edition*. Hampshire, UK: Cengage Learning.

Hava, K., & Ünlü, Z. K. (2021). Investigation of the relationship between middle school students' computational thinking skills and their STEM career interest and attitudes toward inquiry. *Journal of Science Education and Technology*, 1–12. https://doi.org/10.1007/s10956-020-09892-y

Hazari, Z., Potvin, G., Lock, R. M., Lung, F., Sonnert, G., & Sadler, P. M. (2013). Factors that affect the physical science career interest of female students: Testing five common hypotheses. *Physical Review Special Topics-Physics Education Research*, *9*(2), 020115.

Hinckle, M., Rachmatullah, A., Mott, B., Boyer, K. E., Lester, J., & Wiebe, E. (2020, June). The relationship of gender, experiential, and psychological factors to achievement in computer science. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education* (pp. 225–231). https://doi.org/10.1145/3341525.3387403

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, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

IBM Corp. (2017). *IBM SPSS Statistics for Windows, Version 25.0* [Computer Software]. Armonk, NY: IBM Corp.

Irgens, G. A, Dabholkar, S., Bain, C., Woods, P., Hall, K., Swanson, H., Horn, M., & Wilensky, U. (2020). Modeling and measuring high school students' computational thinking practices in science. *Journal of Science Education and Technology*, *29*(1), 137–161. https://doi.org/10.1007/s10956-020-09811-1

Jones, M. G., Corin, E. N., Andre, T., Childers, G. M., & Stevens, V. (2017). Factors contributing to lifelong science learning: Amateur astronomers and birders. *Journal of Research in Science Teaching*, *54*(3), 412–433. https://doi.org/10.1002/tea.21371

Jones, M.G., Chesnutt, K., Ennes, M., Mulvey, K.L., & Cayton, E. (2020). Understanding science career aspirations: Factors predicting future science task value, 1–19. https://doi.org/10.1002/tea.21687

Kier, M. W., & Blanchard, M. R. (2020). Eliciting students' voices through stem career explorations. *International Journal of Science and Mathematics Education*, 1–19. https://doi.org/10.1007/s10763-019-10042-z

Kier, M. W., Blanchard, M. R., Osborne, J. W., & Albert, J. L. (2014). The development of the STEM Career Interest Survey (STEM-CIS). *Research in Science Education*, *44*(3), 461–481. https://doi.org/10.1007/s11165-013-9389-3

Kong, X., Dabney, K. P., & Tai, R. H. (2014). The association between science summer camps and career interest in science and engineering. *International Journal of Science Education, Part B*, *4*(1), 54-65. https://doi.org/10.1080/21548455.2012.760856

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563-575.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*, 79–122. https://doi.org/10.1006/jvbe.1994.1027

Lent, R. W., Brown, S. D., & Hackett, G. (2000). Contextual supports and barriers to career choice: A social cognitive analysis. *Journal of Counseling Psychology*, *47*(1), 36–49. https://doi.org/10.1037/0022-0167.47.1.36.

Lent, R.W., Brown, S. D., & Hackett, G. (2002). Social cognitive career theory. In D. Brown, *Career choice and development fourth edition* (pp. 255-311). San Francisco, CA: Jossey-Bass.

Leonard, J., Buss, A., Gamboa, R., Mitchell, M., Fashola, O. S., Hubert, T., & Almughyirah, S. (2016). Using robotics and game design to enhance children's self-efficacy, stem attitudes, and computational thinking skills. *Journal of Science Education and Technology*, *25*(6), 860–876. https://doi.org/10.1007/s10956-016-9628-2

Linacre, J. M. (2012). *A user's guide to Winsteps [User's manual and software]*. http://www.winsteps.com/winsteps.htm

Loyalka, P., Liu, O. L., Li, G., Chirikov, I., Kardanova, E., Gu, L., ... & Hu, S. (2019). Computer science skills across China, India, Russia, and the United States. *Proceedings of the National Academy of Sciences*, *116*(14), 6732–6736. https://doi.org/10.1073/pnas.1814646116

Lucas, D. M., & Fugitt, J. (2007). The Perception of Math and Math Education in the Rural Midwest. Working Paper No. 37. *Appalachian Collaborative Center for Learning, Assessment, and Instruction in Mathematics (ACCLAIM)*.

Lytle, N., Cateté, V., Boulden, D., Dong, Y., Houchins, J., Milliken, A., ... & Barnes, T. (2019, July). Use, Modify, Create: Comparing computational thinking lesson progressions for STEM classes. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education* (pp. 395–401).

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84

Malcom, S., & Feder, M. (2016). *Barriers and opportunities for 2-year and 4-year STEM degrees: Systemic change to support students' diverse pathways*. Washington, DC: The National Academies Press.

Melguizo, T., & Wolniak, G. C. (2012). The earnings benefits of majoring in STEM fields among high achieving minority students. *Research in Higher Education*, *53*(4), 383–405. https://doi.org/10.1007/s11162-011-9238-z

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741–749.

Moore, T. J., Brophy, S. P., Tank, K. M., Lopez, R. D., Johnston, A. C., Hynes, M. M., & Gajdzik, E. (2020). Multiple representations in computational thinking tasks: A clinical study of second-grade students. *Journal of Science Education and Technology*, *29*(1), 19–34. https://doi.org/10.1007/s10956-020-09812-0

National Academies of Sciences, Engineering, and Medicine. (2018). *Assessing and responding to the growth of computer science undergraduate enrollments*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24926.

National Academies of Sciences, Engineering, and Medicine. (2020). *Promising practices for addressing the underrepresentation of women in science, engineering, and medicine*. Washington, DC: The National Academic Press.

Navlakha, S., & Bar-Joseph, Z. (2011). Algorithms in nature: The convergence of systems biology and computational thinking. *Molecular Systems Biology*, *7*(546), 1–11. https://doi.org/10.1038/msb.2011.78

NGSS Lead States. (2013). *Next generation science standards. For states, by states*. Washington, DC: The National Academies Press.

O'Brien, V., Martinez-Pons, M., & Kopala, M. (1999). Mathematics self-efficacy, ethnic identity, gender, and career interests related to mathematics and science. *The Journal of Educational Research*, *92*(4), 231–235. https://doi.org/10.1080/00220679909597600

Ong, K. J., Chou, Y. C., Yang, D. Y., & Lin, C. C. (2020). Creative drama in science education: The effects on situational interest, career interest, and science-related attitudes of science majors and non-science majors. *EURASIA Journal of Mathematics, Science and Technology Education*, *16*(4), 1–18. https://doi.org/10.29333/ejmste/115296

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*(9), 1049–1079. https://doi.org/10.1080/0950069032000032199

Pallant, A., & Lee, H. S. (2015). Constructing scientific arguments using evidence from dynamic computational climate models. *Journal of Science Education and Technology*, *24*(2–3), 378–395.

Papadimitriou, C. H. (2003). MythematiCS: in praise of storytelling in the teaching of computer science and math. *ACM SIGCSE Bulletin*, *35*(4), 7–9.

Peel, A., Sadler, T. D., & Friedrichsen, P. (2019). Learning natural selection through computational thinking: Unplugged design of algorithmic explanations. *Journal of Research in Science Teaching*, *56*(7), 983–1007. https://doi.org/10.1002/tea.21545

Pierson, A. E., & Clark, D. B. (2018). Engaging students in computational modeling: The role of an external audience in shaping conceptual learning, model quality, and classroom discourse. *Science Education*, *102*(6), 1336–1362. https://doi.org/10.1002/sce.21476

Pierson, A. E., Brady, C. E., & Clark, D. B. (2020). Balancing the environment: Computational models as interactive participants in a STEM classroom. *Journal of Science Education and Technology*, *29*(1), 101–119.

Price, C. A., Kares, F., Segovia, G., & Loyd, A. B. (2019). Staff matter: Gender differences in science, technology, engineering or math (STEM) career interest development in adolescent youth. *Applied Developmental Science*, *23*(3), 239–254. https://doi.org/10.1080/10888691.2017.1398090

Rachmatullah, A., Mayhorn, C. B., & Wiebe, E. N. (2021). The effects of prior experience and gender on middle school students' computer science learning and monitoring accuracy in the Use-Modify-Create progression. *Learning and Individual Differences*, *86*, 101983. https://doi.org/10.1016/j.lindif.2021.101983

Rachmatullah, A., Roshayanti, F., Shin, S., Lee, J. K., & Ha, M. (2018). The secondary-student science learning motivation in Korea and Indonesia. *EURASIA Journal of Mathematics, Science and Technology Education*, *14*(7), 3123–3141. https://doi.org/10.29333/ejmste/91665

Rachmatullah, A., Wiebe, E., Boulden, D., Mott, B., Boyer, K., & Lester, J. (2020). Development and validation of the computer science attitudes scale for middle school students (MG-CS attitudes). *Computers in Human Behavior Reports*, *2*, 1–11. https://doi.org/10.1016/j.chbr.2020.100018

Raubenheimer, J. (2004). An item selection procedure to maximize scale reliability and validity. *SA Journal of Industrial Psychology*, *30*(4), 59–64. https://hdl.handle.net/10520/EJC89023

Revelle, W. (2018). *psych: Procedures for personality and psychological research*. Evanston, IL: Northwestern University.

Roller, S. A., Lampley, S. A., Dillihunt, M. L., Benfield, M. P., Gholston, S. E., Turner, M. W., & Davis, A. M. (2020). Development and initial validation of the Student Interest and Choice in STEM (SIC-STEM) survey 2.0 instrument for assessment of the social cognitive career theory constructs. *Journal of Science Education and Technology*, *29*(5), 646–657. https://doi.org/10.1007/s10956-020-09843-7

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

Sáinz, M., & Eccles, J. (2012). Self-concept of computer and math ability: Gender implications across time and within ICT studies. *Journal of Vocational Behavior*, *80*(2), 486–499. https://doi.org/10.1016/j.jvb.2011.08.005

Salac, J., Thomas, C., Twarek, B., Marsland, W., & Franklin, D. (2020, February). Comprehending code: Understanding the relationship between reading and math proficiency, and 4th-grade CS learning outcomes. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (pp. 268–274). https://doi.org/10.1145/3328778.3366822

Saw, G., Chang, C. N., & Chan, H. Y. (2018). Cross-sectional and longitudinal disparities in STEM career aspirations at the intersection of gender, race/ethnicity, and socioeconomic status. *Educational Researcher*, *47*(8), 525–531. https://doi.org/10.3102/0013189X18787818

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, *29*(4), 304–321. https://doi.org/10.1177/0734282911406653

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, *99*(6), 323–338.

Schwab, K., & Sala-i-Martín, X. (2012). *Insight report: The global competitiveness report 2012–2013*. Geneva: World Economic Forum.

Settlage, J., & Meadows, L. (2002). Standards-based reform and its unintended consequences: Implications for science education within America's urban schools. *Journal of Research in Science Teaching*, *39*(2), 114–127. https://doi.org/10.1002/tea.10012

Shapiro, H., Østergaard, S. F., & Hougaard K. F. (2015). *Does Europe need more STEM graduates?: Final report*. Luxembourg: Publications Office of the European Union.

Shin, S., Rachmatullah, A., Roshayanti, F., Ha, M., & Lee, J. K. (2018). Career motivation of secondary students in STEM: A cross-cultural study between Korea and Indonesia. *International Journal for Educational and Vocational Guidance*, *18*(2), 203–231. https://doi.org/10.1007/s10775-017-9355-0

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

Skamp, K. (2007). Conceptual learning in the primary and middle years: The interplay of heads, hearts, and hands-on science. *Teaching Science*, *53*(3), 18–22.

Smith, R. M. (1996). A comparison of methods for determining dimensionality in Rasch measurement. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(1), 25–40. https://doi.org/10.1080/10705519609540027

Smith, S. M. (2002). The role of social cognitive career theory in information technology based academic performance. *Information Technology, Learning, and Performance Journal*, *20*(2), 1–10.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, UK: Cambridge University Press.

Turner, S. L., Joeng, J. R., Sims, M. D., Dade, S. N., & Reid, M. F. (2019). SES, gender, and STEM career interests, goals, and actions: A test of SCCT. *Journal of Career Assessment*, *27*(1), 134-150. https://doi.org/10.1177/1069072717748665

U.S. Bureau of Labor Statistics. (2017). *STEM occupations: Past, present, and future*. Retrieved from https://www.bls.gov/spotlight/2017/science-technology-engineering-and-mathematics-stem-occupations-past-present-and-future/home.htm.

Unfried, A., Faber, M., Stanhope, D. S., & Wiebe, E. (2015). The development and validation of a measure of student attitudes toward science, technology, engineering, and math (S-STEM). *Journal of Psychoeducational Assessment*, *33*(7), 622–639. https://doi.org/10.1177/0734282915571160

Wagh, A., Cook-Whitt, K., & Wilensky, U. (2017). Bridging inquiry-based science and constructionism: Exploring the alignment between students tinkering with code of computational models and goals of inquiry. *Journal of Research in Science Teaching*, *54*(5), 615–641. https://doi.org/10.1002/tea.21379

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

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

Wiebe, E., Unfried, A., & Faber, M. (2018). The relationship of STEM attitudes and career interest. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(10), 1–17. https://doi.org/10.29333/ejmste/92286

Wilkerson-Jerde, M. H., Gravel, B. E., & Macrander, C. A. (2015). Exploring shifts in middle school learners' modeling activity while generating drawings, animations, and computational simulations of molecular diffusion. *Journal of Science Education and Technology*, *24*(2–3), 396–415. https://doi.org/10.1007/s10956-014-9497-5

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean square fit values. *Rasch Measurement Transactions*, *8*(3), 370.

CHAPTER 4: Developmental Trends and Sources of Middle School Teachers' Self-efficacy for Teaching Science in a Computationally Rich Environment: A Mixed-Methods Study

Abstract

The inclusion of computational thinking (CT) into science curricula has advocated implementing a computationally rich science learning environment where students learn science via building models in a computer programming platform. Such an approach may influence teachers' self-efficacy for teaching science which may also be associated with their self-efficacy for teaching CT. Framed using Bandura's Social Cognitive Theory, this study investigated the development and sources of teachers' self-efficacy for teaching science and CT and looking at whether the two constructs are related to one another. A total of eleven middle school science teachers (seven in-service and four pre-service) participated in a CT-integrated science instruction workshop. They then implemented the curriculum they learned and obtained from the workshop in their classrooms. The teachers took questionnaires on science and CT teaching efficacy beliefs four times: before and after the workshop and before and after they taught. As a follow-up, interviews and writing reflections were collected after they took the instruments. Skillings-Mack and repeated-measures correlation tests were run on the quantitative data, and the qualitative data were analyzed thematically. Results from quantitative analyses revealed a pattern of increasing teachers' self-efficacy for teaching science and CT in a computationally rich environment over the administrations of the instrument. Thematic analysis showed three sources of teachers' self-efficacy: computer programming experience, students' interests, and teaching repetition and field experience. This study calls attention to the importance of providing experience for teachers to teach science in a computationally rich environment, whether through professional development or teacher education programs.

Keywords: computational thinking, mixed-methods, middle school, science teacher, teaching efficacy beliefs

Introduction

The growing and accelerating number of computationally-intensive Science, Technology, Engineering, and Mathematics (STEM) jobs has supported the importance of nurturing computational thinking (CT) in K–12 education (Google & Gallup, 2016; Grover & Pea, 2018; National Research Council, 2012). Although consensus has not yet been reached in the literature on the precise definition of CT, many have consistently argued that CT is a set of skills related to solving computational and complex problems by using computer science (CS) concepts, tools, and principles. However, experts have argued that CT does not necessarily involve computing devices; CT can also be used when solving problems without computers (for a review, see Shute et al., 2017; Tang et al., 2020). Wing (2006) contended that CT skills are crucial for students to prepare for employment and an everyday life full of digital and computational products, services, and problems. To address these issues, educators have integrated CT into K–12 curriculum and standards outside of CS-specific courses. One area to integrate CT is science: for example, the Next Generation Science Standards (NGSS Lead States, 2013) lists CT as among the eight essential science and engineering practices. Consequently, the inclusion of CT in NGSS has proliferated a line of research on computationally rich science activities where students build a computational model of a scientific phenomenon via computer programming (Lee & Malyn-Smith, 2019; Weintrop et al., 2016). Research has suggested that such integration can advance both students' science conceptual understanding and computational thinking skills (Aksit & Wiebe, 2020; Irgens et al., 2020; Peel et al., 2019), as well as their attitudes toward science and computing (Leonard et al., 2016; Wagh et al., 2017).

In creating a computational model, students are expected to understand both scientific concepts and CS concepts, as they are used within computer programming environments. In addition to what students are expected to know and do, science teachers must grasp the material at a level that allows them to teach and facilitate students' learning of both science and CS concepts. The added burden of CS concepts—and programming—presents a challenge because teachers often lack expertise in these subjects, which are rarely taught in their education programs (Langbeheim et al., 2020; Rich et al., 2021). Such burdens may also affect teachers' self-efficacy for teaching science in a computationally rich science environment. According to d'Alessio's (2018) study on preservice teachers' teaching efficacy beliefs from the lens of Bandura's social cognitive theory, content mastery is among the significant sources of positive teaching-efficacy beliefs. As many scholars have noted, teachers' teaching self-efficacy is an important factor in determining students' learning outcomes (e.g., Althauser, 2015; McCaffrey et al., 2003; Tschannen-Moran & Barr, 2004; Zee & Koomen, 2016). When teachers lack confidence in teaching science in a computationally rich environment, students may not optimally develop their conceptual understanding of science and CT skills. Hence, understanding how science teachers develop their self-efficacy for teaching science in a CT-integrated environment, as well as the sources of such development, are imperative to help elevate teachers' self-efficacy and, in the process, maximizing students' potential in the learning of both science concepts and CT skills.

Researchers have begun investigating the development of science teachers' self-efficacy for teaching in computationally rich environments. These studies have used either quantitative or qualitative, pre–post design, and have focused largely on elementary teachers. For instance, Jaipal-Jamani and Angeli (2017) quantitatively examined the impact of robotics on elementary

preservice teachers' self-efficacy, programming skills, and science conceptual understanding. They found an increase in the preservice teachers' programming ability and science conceptual understanding, as well as their self-efficacy. In addition, Ketelhut et al. (2020) recently conducted a qualitative study on elementary teachers' changing beliefs about integrating CT into elementary science after they engaged in a 2-day professional development (PD). Although the participating teachers were still unsure about finding resources and integrating CT into elementary science instruction, the authors found the teachers developed a positive view about infusing CT into science instruction and even felt motivated to design CT-based science instruction.

This initial work has provided valuable insights into how PD and intervention may change teachers' beliefs about CT-integrated science learning. However, these studies did not further leverage quantitative and qualitative data analyses to provide evidence for the interventions' effectiveness. Langbeheim et al. (2020) conducted a mixed-methods study on middle school physics teachers after a computational modeling workshop. They found that teachers' prior experience in programming significantly influenced their science teaching in computationally rich environments. However, Langbeheim and colleagues did not delve into the subject to uncover other teaching-efficacy sources. Nor did they follow up on the state of teachers' teaching efficacy after implementing their new-found knowledge and skills in the classroom. According to Bandura (1977), this teaching experience is critical given that it is part of mastery experience which is among the pivotal sources of self-efficacy. To more fully understand how teacher efficacy in CT-integrated science progresses as a result of both PD and teaching, in the current study, we investigated teachers by following their self-efficacy development for teaching science and CT before and after a short workshop on infusing CT into

science instruction. In addition, the study's duration covered the time following the teachers' CT incorporation into the classrooms.

This study examines the changes in teachers' science and CT teaching-efficacy beliefs and explores sources of such changes. We believe that this study's findings may contribute to the fuller understanding of teachers' developmental trends in teaching-efficacy beliefs in an integrated environment, particularly CT-integrated science instruction. The following research questions guided the current study:

1. What impacts do a short CT-integrated science teaching workshop and classroom teaching implementation have on the developmental trends of middle school teachers' self-efficacy for teaching science and CT?
2. To what extent does development of self-efficacy for teaching science in a computationally rich science activity associated with self-efficacy changes for teaching CT?
3. What are the contributing sources to the development of teachers' science- and CT-teaching efficacy?

Theoretical Framework: Social Cognitive Theory

The current study is framed by Bandura's (1977) social cognitive theory (SCT). SCT describes how people obtain and maintain certain decisions contributing to their behavioral patterns. According to SCT, human functioning is processed via reciprocal causation of three determinants: behavior, environmental events, and personal factors (including cognitive and biological aspects). Pajares (1997) argued about the possibilities for change by targeting one of the three determinants. Consistent with this, Bandura (1997) also believed that self-efficacy, one of the core components of SCT, is susceptible to change under specific conditions.

Self-efficacy and outcome expectations are at the heart of SCT. Tschannen-Moran and Hoy (2001) described teacher self-efficacy (TSE) as "judgment about [teacher's] ability to achieve the desired results of students' engagement and learning, even among students who may be difficult or unmotivated" (p. 783). TSE primarily focuses on teachers' beliefs about their abilities to perform certain tasks within the realm of teaching and learning. Lakshmanan et al. (2011) explained that teachers' outcome expectations (beliefs about a particular behavior, e.g., teaching methods) would give rise to a specific outcome. Williams et al. (2005) noted that outcome expectations emerge from self-efficacy and directly control and influence behavior. Gibson and Dembo (1984) and Lauermann and Karabenick (2011) further added that teachers' outcome expectations are influenced by external factors such as students' interests, limiting the teacher's ability to impact students' optimal learning. Teacher education research literature has not yet reached a consensus about measuring teachers' outcome expectations. The measure should be tied to teachers' actions (Bandura, 2006). However, the instruments available in science teacher education, such as STEBI (Riggs & Enoch, 1990) or its refinement version, such as T-STEM Science Scale (Author et al., under review), measure more general expectancy beliefs referring to Rotter's (1966) locus of control and teachers' responsibility. Locus of control is conceptualized as a type of general expectation that connects to students' learning but does not require a close connection between teachers' actions and students' outcomes (Dellinger et al., 2008). Due to this lack of clarity in measurement definition, we focused more on TSE in this study.

Literature has indicated that there are four primary sources of influence on teachers' self-efficacy: (a) past or mastery experiences, (b) vicarious experiences, (c) social persuasion, and (d) physiological and emotional arousals (Bandura, 1977; Bong & Skaalvik, 2003). Mastery

experiences are related to teachers' experiences in mastering or having prior learning of the content knowledge when they were students or even their experiences in teaching the content knowledge (d'Alessio, 2018; Menon & Sadler, 2016; Langbeheim et al., 2020). In the context of this study, these experiences may include their content mastery of both science and CT or CS concepts. These experiences are imperative, especially as they tend to shape teachers' attitudes toward the content they teach (Jong & Hodges, 2013). The second source, vicarious experience, can be considered teachers' experiences in successfully delivering science concepts with a certain teaching approach. Vicarious experience can also be obtained from observing an expert or exemplary teaching and learning models. Next, social persuasion is defined as the feedback teachers obtain from peers, administrators, or mentors (Tschannen-Moran & McMaster, 2009). Bandura (1977) indicated that the persuaders' credibility and expertise are highly tied to how teachers may alter their teaching-efficacy beliefs. Both vicarious experiences and social persuasion are highly relevant to teachers' PD, workshops, or educational training experiences (Rupp & Becker, 2021; Wei, 2020; Zangori et al., 2018). Last, teachers' self-efficacy can also be influenced by their psychological or emotional states, such as excitement or anxiety (Thomson et al., 2019).

Literature Review

Computational Thinking and Science Education

Research has indicated that CT can be viewed from two distinct yet related perspectives: CS contexts and general problem-solving contexts (see Tang et al., 2020). Following Langbeheim et al.'s (2020) approach, the current study acknowledges both perspectives by considering the ideas of CT as problem-solving skills that can be acquired through computer programming activities. Based on this, Wing's (2008) definition of CT is still relevant to the

current study. Wing defined CT as “an approach to solving problems, designing systems, and understanding human behavior that draws on concepts fundamental to computing” (p. 3).

Integrating CT into K–12 education is beneficial in this digital era where individuals should have advanced technological and computational literacies to enable informed and active decision-making (Wing, 2006, 2008). Swanson et al. (2019) also argued that incorporating CT into K–12 education by leveraging computer science logic may allow for a shift in conceptualizing students from technology users to creators. Furthermore, learning CT through computer programming and infusing this approach into science instruction has been found to help students learn both science concepts and practices (Aksit & Wiebe, 2020; Irgens et al., 2020; Peel et al., 2019).

Although research and literature have indicated that integrating CT into science instruction through programming holds great potential for students’ science learning, this strategy requires teachers to obtain a certain level of programming fluency (Langbeheim et al., 2020; Peel et al., 2019). As mentioned previously, teachers’ mastery experiences significantly contribute to their teaching-efficacy beliefs. Thus, adding computer programming fluency to science instruction may influence the dynamic of teachers’ teaching-efficacy beliefs in a computationally rich science environment. However, most previous studies on teachers’ teaching efficacy in CT-CS or technology-integrated science instruction have still not explored to what degree this dynamic association appears (e.g., Adler & Kim, 2018; Ketelhut et al., 2020; Langbeheim et al., 2020). The following section provides a review of this topic.

Developmental Trends of Teachers' Self-efficacy for CT-CS or Technology-Integrated Science Instructions

Integrating technologies, including but not limited to CT and CS, into science instruction has been considered a component of reformed-based science instruction (Menon et al., 2020). Hence, experts have attempted to develop professional learning opportunities for preservice and in-service teachers to improve their science teaching efficacy for using technologies (e.g., Adler and Kim, 2018; Langbeheim, 2020). For instance, Campbell et al. (2015) explored science teachers' tendencies to implement reformed-based science instruction by integrating computer technology into their teaching practices. The participating teachers engaged in a yearlong professional development (PD) on the instructional technologies. The authors found a significant improvement in teachers' reformed-based science teaching and the use of technology scores after the PD. In addition, the teachers also learned new information and communications technology (ICT) skills from participating in the PD.

Building on the above literature and transferring the implications to computer programming or CT-CS areas, Hadjiachilleos et al. (2013) introduced Lego programming to preservice teachers to help increase their engagement in scientific inquiries. Through a qualitative analysis of the teachers' interactions, the authors found that teachers were more involved in scientific inquiries while using the Lego environment. Adler and Kim (2018) added to these results by taking a more progressive approach of leveraging the Hour of Code tutorial combined with web-based simulation in a science methods course to provide CT understanding for preservice science teachers. Their interview analyses revealed that teachers understood the physics content and they felt that using programming software would help them integrate CT into their science instruction. Langbeheim et al.'s (2020) findings differed from Adler and Kim's

study when examining the impact of a 30-hour computational modeling physics workshop on 9th-grade in-service science teachers. They focused on investigating the change in the teachers' physics teaching self-efficacy and programming confidence and did not find any significant changes in the participating teachers' physics teaching efficacy. However, when teachers' years of experience were considered (novice or experienced teachers), they found that experienced teachers had significantly higher scores on this construct. Moreover, the researchers found a significant increase in teachers' programming self-efficacy from before to after the workshop. Framed by Clarke and Hollingsworth's (2002) interconnected model of professional growth (IMPG), Ketelhut et al. (2020) added to this finding by conducting a qualitative study on mentor and preservice elementary science teachers after they participated in a two-day workshop about CT integration in elementary science learning. The PD workshop focused primarily on educating the teachers about Weintrop et al.'s (2016) CT taxonomy, contextualized for elementary science instruction. After the PD workshop, the authors found the participating teachers were positive about integrating CT into their teaching, believed that CT could offer the best science teaching practices, and even implemented it in their science instruction when realizing how the integrated activities could highly engage the students. However, the authors' analysis of written reflections and interviews also indicated that no computational integration was found in the self-reported practices, and the teachers' students struggled to engage in CT's higher-level thinking. Furthermore, the participating teachers also reflected that time and resources hindered their attempts to implement CT-integrated science learning in their classrooms.

Although most of the above-mentioned studies have provided valuable insights into the impact of professional learning opportunities on teachers' teaching efficacy in science and CT-related technologies, most studies did not further examine teachers after they implemented the

content and strategies from the PD in their classrooms. Thomson et al. (2019) and d'Alassio (2018), who each conducted a longitudinal study with preservice science teachers after they taught in a microteaching setting, found that teaching experiences significantly influenced their teaching efficacy favorably. Additionally, the previous cited studies on CT and technology-integrated science instruction only explored science and CT-CS or technology efficacy separately. However, they may share a reciprocal relationship in such an integrated learning environment. This relationship is evident in Wei's (2020) study, which explored the development of Chinese high school science teachers' perceptions and implementation of integrated science teaching and learning. Wei found that teachers actively reshaped their science teaching conception given that, in integrated science instruction, teachers teach more than only science concepts. Therefore, in the current study, we aimed to fill this literature gap by examining the developmental trends of teachers' self-efficacy for teaching science and CT and whether science teaching efficacy in a computationally rich environment interacts with CT teaching efficacy.

Methods

Research Design

The current study employed a mixed-methods explanatory sequential design (Creswell & Creswell, 2018). The explanatory sequential design consists of two phases: quantitative components in the first and qualitative in the second. Applying this two-phase analysis cycle to the study design resulted in four iterations: before the workshop, after the workshop, before teaching, and after teaching. Instead of selecting specific participants in the qualitative phase, as Creswell and Creswell (2018) suggested, we involved all our participants in both phases because our sample size was limited. We believed that this approach would enable us to better understand the development of teachers' CT and science teaching efficacy.

Participants and Context

A total of 11 Indonesian middle school teachers participated in this study. Six were in-service female teachers, one in-service male, two pre-service female teachers, and two pre-service male teachers. The in-service teachers' teaching experiences ranged from two to 32 years. The four preservice science teachers were in their last semester of a prominent public teacher college's teacher preparation program. Teaching in a school for about six months is required for students who are enrolled in an Indonesian teacher college and have already completed the prerequisite courses. They usually teach their subject expertise in private or public schools, mentored by a college professor and a mentor teacher. Almost all the participating teachers had a background in biology or biology education, except Regina, who had physics experience, and Aiden (all names are pseudonyms), who had a CS-education background. Regardless of their disciplinary backgrounds, all in-service teachers were qualified to teach science in the middle grades, as middle school science covers more general science topics. Table 1 provides more detailed information about the 11 teachers.

All the teachers participated in a 5-hour long online training workshop delivered by one of the authors on integrating CT into science instruction. This workshop focused on an instructional strategy for CT-CS called the Use-Modify-Create (UMC) progression developed by Lee et al. (2011). During this workshop, the teachers were introduced to building and teaching a computational model of the food web. Our research team developed the activities which underwent several rounds of modifications framed by design-based research (Lytle et al., 2019; Rachmatullah et al., 2021). Together with the workshop instructor, the teacher reviewed all the computational modeling activities, from unplugged (without computer programming) to each stage in the UMC progression. In the *unplugged activity*, the teachers were engaged in lecture-

and drawing-based instructions to learn the core concepts of the food web and computational modeling through drawing the food webs and writing pseudocode (i.e., notations resembling simplified programming language). They were told that this activity's primary purpose was to introduce the core scientific concepts to students and help them plan their computational models using pseudocode.

Table 1

Demographics Information of the Participants

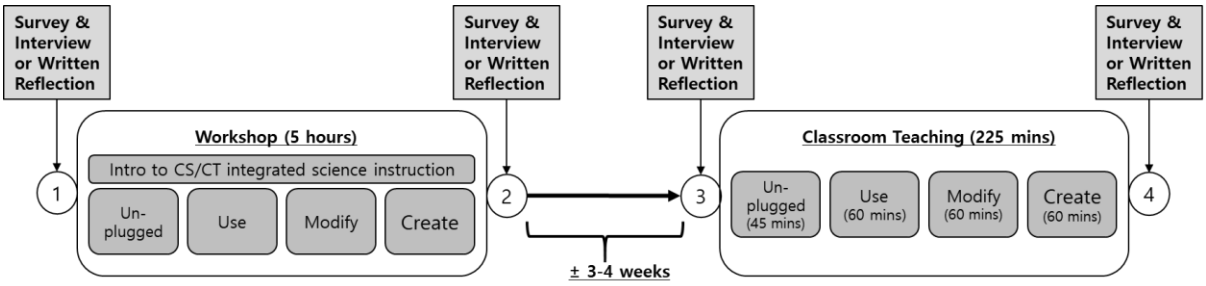
Teacher (Pseudonym)	Teacher Status	Gender	Year(s) of Experience	School	Subject
Wilma	In-service	Female	18 years	Private	Biology
Regina	In-service	Female	32 years	Public	Physics
Dina	In-service	Female	5 years	Private	Biology
Shirley	In-service	Female	5 years	Private	Biology
Irene	In-service	Female	17 years	Public	Biology
Alisha	In-service	Female	2 years	Private	Biology
Aiden	In-service	Male	3 years	Private	Computer Science
Tamara	Pre-service	Female	0 year	Public	Biology
Avery	Pre-service	Male	0 year	Private	Biology
Damian	Pre-service	Male	0 year	Public	Biology
Vivian	Pre-service	Female	0 year	Public	Biology

The second activity is the *Use* activity. The Use activity engaged teachers with using pre-existing computational model visible code that demonstrates the relationship between solar energy and plant growth. This activity was followed by a *Modify* activity where the teachers added new code to create a new actor (bunny) to their food web model. They also modified the plants' code according to the plants' responses to the bunnies eating them. For the bunny code, the teachers were given an incomplete and wrong set of code blocks to debug. Last, the *Create* activity engaged teachers with building the computational code from scratch to model the secondary consumer (fox). In all of these activities, the teachers used a worksheet to collect data

from their computational model and the reasons for their gathered data. After they finished the workshop, we provided them with a curriculum guide for the activity they had just completed and asked them to implement it in their classrooms. On average, three weeks after the workshop, the teachers started implementing the classroom activity for about four consecutive days (class periods). Each day served one activity that lasted for 60 minutes and was done remotely and synchronously via Zoom due to the pandemic. Figure 1 shows the sequence of the activity and data collection process.

Figure 1

Data Collection Procedure



Notes: 1 = pretraining, 2 = posttraining, 3 = preteaching, 4 = postteaching

Data Collection and Sources

The data sources included both quantitative survey instruments and qualitative data through interviews or teachers' written reflections. As depicted in Figure 1, the data sources came from four different times: both before and after the workshop and classroom implementation. We collected data after the workshop and before the classroom teaching, because, as Adler and Kim (2018) suggested, teachers expressed their intention to familiarize

themselves with the programming platform and obtain more practice during this time period. We believed that collecting data before the classroom teaching would provide additional insights into teachers' CT and science teaching efficacy beliefs and mediate any history threat to validity (Campbell & Stanley, 1963).

Self-report Instruments

Two instruments measuring teachers' CT and science teaching efficacy beliefs were utilized during the four-time points. The same version of the instruments was used each time. The T-STEM Science scale developed by Unfried et al. (under review) was administered to gather data on teachers' self-efficacy for teaching science. The T-STEM Science scale is a refinement of the Science Teaching Efficacy Belief Instrument A (STEBI-A, Riggs & Enochs, 1990), and was developed in response to perceived shortcomings of the STEBI-A identified in the literature (e.g., Betebenner, 2009; Ho, 2008; Lachlan-Hache & Castro, 2015). The T-STEM Science scale—self-efficacy subscale— consisted of nine items. An example item from this instrument is, "I am confident that I can teach science effectively." To measure teachers' self-efficacy for teaching CT, we used the T-STEM CT scale developed and validated by Boulden et al. (2021). Before taking the T-STEM CT, especially during the first data collection, all teachers were provided with a written definition of CT to help minimize differences in the interpreted definition of the primary construct. An example item in this instrument includes, "I know the steps necessary to teach computational thinking effectively." All items in the two chosen instruments were 5-point Likert questions (1 = *strongly disagree* to 5 = *strongly agree*). Given all the instruments were initially developed in English, we followed the back-translation procedure suggested by Brislin (1970) and Cha et al. (2007) to translate them into Bahasa Indonesia.

Semi-Structured Interview and Written Reflection

In addition to the self-report instruments, the teachers were asked to be interviewed and provide a written reflection on their confidence in teaching science in a computationally rich environment. Both semi-structured interviews and written reflections were intended to explain the states of teachers' teaching efficacy at each time point. The semi-structured interview duration was, on average, between 15–30 minutes. The following are examples of questions posed to the teachers during the interview sessions:

- On what do you base your answers to these statements on both Science and CT?
- Can you describe any teaching areas that you feel more or less confident in and why?
- What efforts do you think you can employ to help maintain or improve your confidence in teaching science?
- Can you describe your confidence in teaching about the food web in such environments? Why do you feel this way?
- What obstacles do you think you would face when teaching this to students?
- Can you describe your feelings about today's teaching?
- What do you think was the most challenging aspect?

Data Analyses

Quantitative

A non-parametric Skillings-Mack test was performed to answer the first research question regarding the development of teachers' self-efficacy for teaching science and CT. A non-parametric test was selected because of the small sample size (< 30) which is highly likely to be not normally distributed (Agresti, 2018). The Skillings-Mack test is equivalent to Friedman's test and is used when some data are missing (Srisuradetchai & Trakultraipruk, 2016). In our

sample, two teachers, Irene and Damian, did not continue teaching after the workshop due to logistical issues (e.g., too few students possessing a computer); thus, we lacked a complete dataset for them. A posthoc test using Wilcoxon signed-rank test with Bonferroni correction followed the Skillings-Mack test to examine further the specific point where the change was statistically significant. In this posthoc, given that we had four time-points, Bonferroni correction would move the significance level to alpha = .05/4 = .0125. Thus, we set this alpha value in the posthoc, meaning that a significant difference would occur where the alpha value is less than or equal to .0125. The Skillings-Mack test was run in R, and the Wilcoxon signed-rank test was performed in IBM SPSS V26.

Additionally, a repeated measure correlation test was performed to answer the second research question regarding the relationship between CT and science teaching efficacy beliefs. Bakdash and Marusich (2017) explained that this test is a "statistical technique for determining the common within individual association for paired measures assessed on two or more occasions for multiple individuals" (p. 1). A repeated measure correlation test was run using the rmcorr package in RStudio.

Qualitative

Data collected through semi-structured interviews and written reflection were analyzed qualitatively to answer the third research question regarding sources of teachers' science and CT teaching efficacy at each development stage. A constant comparative analysis method was used to detect emerging patterns from the participating teachers (Strauss & Corbin, 1998). Three phases of the qualitative coding processes were employed. The first phase was an open coding process, mostly sentence-by-sentence, to gather each segmented data's meaning. Next was axial coding, where all the codes generated (i.e., sub-categories) from the open coding process were

grouped into more abstract concepts (i.e., categories). These categories were primarily the four sources of self-efficacy defined by SCT: (a) past or mastery experiences, (b) vicarious experiences, (c) social persuasion, and (d) physiological and emotional arousals. Sub-categories were generated and grouped into the four sources. A codebook (Table 2) was established and given to a trained second coder along with the coding rules. The second coder coded 20% of the segmented data to generate an inter-coder reliability statistic (*k*). The final *k* was .937, indicating satisfactory agreement. Any disagreement was discussed until consensus was reached, which might have led to some modifications of the codebook. The last phase was selective coding, where we chose a category/sub-category as a focus to generate themes and a constructive story based upon the connections of that category/sub-category with other categories/sub-categories (Saldaña, 2013). The emerging themes were shared and discussed with the second coder through peer debriefing to confirm them (Patton, 2002). Furthermore, both coders used memos to record any methodological decisions and initial thoughts on emerging themes throughout the coding process.

It is noteworthy to emphasize that all the written reflections and interviews were conducted in Indonesian and were translated to English for reporting purposes. The translation was done as accurate as possible to the original semantic intent, even if a typical English translation might shift the language more for grammatical correctness. We were also aware that the translation might not precisely show the subtle contextual expression of the teachers. Hence, we acknowledge this reporting limitation.

Table 2

Codebook and Inter-Coder Reliability (Cohen's k) for Each Category and Sub-Category

Category	Sub-Category	Description	Example
Past or mastery experience (k = .967)	Students' characteristics (k = 1.000)	Teachers talked about students' characteristics and how their characteristics influenced teaching efficacy beliefs.	"Students also already have some experiences with computer programming from robotics classes" (Alisha, posttraining, Line 85).
	Educational experience and concept mastery (k = 1.000)	Teachers expressed their academic experiences with learning scientific concepts and how they mastered those concepts.	"My background is not in science" (Aiden, preteaching, Line 246). "I think I am confident with my understanding of the scientific concepts and theories" (Shirley, pretraining, Line 7).
	Programming experience (k = .900)	Teachers discussed their prior experience with programming or CS-related activities, as well as their familiarity with CT.	"I understand the fundamental idea of programming language and am familiar with the platform we are about to use" (Dina, posttraining, Line 49).
Vicarious experience (k = .919)	Everyday teaching practices (k = .838)	Teachers expressed that their confidence was based on their day-to-day teaching practices and experiences, including media use.	"I probably would just go to an open field or rice field, so students could directly observe the plants, mice, and snakes" (Damian, pretraining, Line 158).
	Professional development (k = 1.000)	Teachers talked about how participating in professional development, as well as training, influenced their confidence in teaching science and CT.	"After training, I now understand and feel more confident teaching science using coding" (Dina, posttraining reflection, Line 45).

Table 2

Codebook and Inter-Coder Reliability (Cohen's k) for Each Category and Sub-Category (Continued…)

Category	Sub-Category	Description	Example
Vicarious experience (k = .919)	Preparation and repetition (k = .839)	Teachers expressed that repetition, preparation, and practice are the sources of their confidence in teaching science and CT.	"I feel more confident when I have tried it out. Before, I just practiced several times to become more confident" (Alisha, postteaching, Line 10).
	Exemplary teaching models (k = 1.000)	Teachers said that watching other teachers teach helped elevate their confidence.	"I feel confident after observing how other teachers teach" (Avery, preteaching, Line 183).
Social persuasion (k = 1.000)	Colleagues' feedback and comments (k = 1.000)	Teachers stated that feedback and comments from colleagues impacted their confidence in teaching science and CT.	"I also asked my observer for feedback or talked to you about my teaching" (Wilma, postteaching, Line 257).
Physiological and emotional arousals (k = .897)	Students' responses and interests (k = .869)	Teachers expressed that students' responses and interests during the activities influenced their confidence in teaching science and CT.	"The important thing is not students' understanding but their interest. Usually when they [students] are already interested in the activity, it will lift my energy and confidence" (Dina, pretraining, Line 10).
	Equipment, time, and technical problems (k = .924)	Teachers talked about how time allocation, computer availability, internet connectivity, and other technical problems influenced their confidence during teaching and learning activities.	"When students run into internet problems, time allocation is affected" (Wilma, preteaching, Line 79).

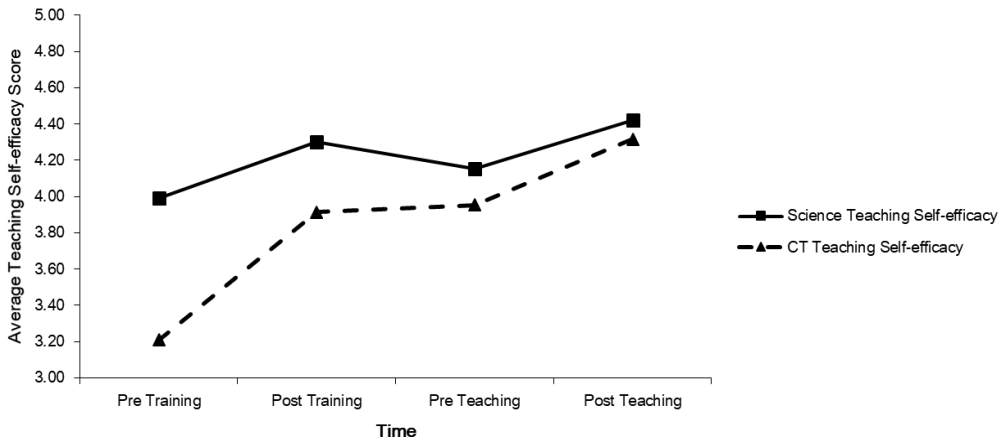
Findings

RQ1: The Development of Teachers' Self-Efficacy for Teaching Science and CT

The Skillings-Mack test results revealed a significant increasing trend in teachers' self-efficacy for teaching science: *Skillings-Mack Statistic* = 12.93, *df* = 3, *p* = .005. Such a significant increasing trend was also seen in teachers' self-efficacy for teaching CT: *Skillings-Mack Statistic* = 12.83, *df* = 3, *p* = .005. The results are visualized in Figure 2.

Figure 2

The Increasing Trends of Teachers' Science and CT Teaching Self-Efficacy



We followed up these results by performing the Wilcoxon signed-rank test with Bonferroni correction ($p < .0125$) to examine when the significant changes occurred. For science teaching self-efficacy, we did not find a significant increase from pretraining to posttraining ($Z = -2.26$, $p = .024$), posttraining to preteaching ($Z = -0.54$, $p = .589$), and preteaching to postteaching ($Z = -1.88$, $p = .061$). Similarly, for CT teaching self-efficacy, results revealed non-significant changes from pretraining to posttraining ($Z = -2.24$, $p = .025$), posttraining to preteaching ($Z = -0.85$, $p = .933$), and preteaching to postteaching ($Z = -1.19$, $p = .233$). In

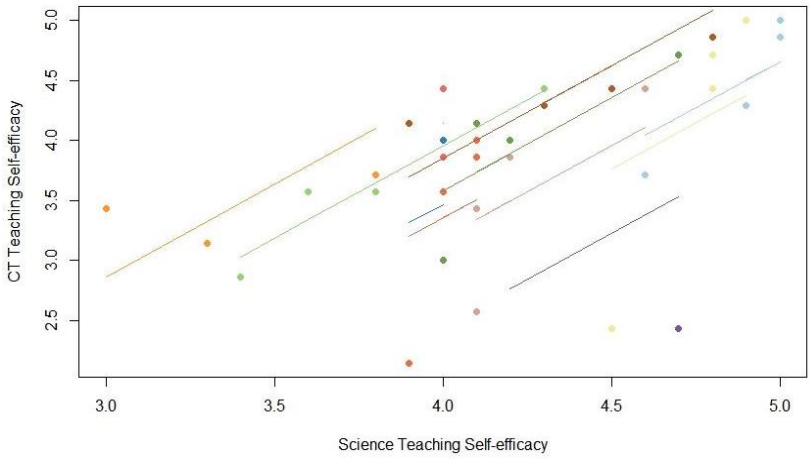
addition, we found a closing pattern of the gap between self-efficacy for teaching science and CT from pretraining to postteaching. Interestingly, both self-efficacy for teaching science and CT continued to rise after teaching, which might indicate teachers' understanding of the value of teaching science in a computationally rich environment.

RQ2: The Relationship between Self-efficacy for Teaching Science and CT

A repeated-measures correlation test was run to examine the relationship between teachers' self-efficacy for teaching science and CT using repeated-measures data. We found a significant medium-high correlation between self-efficacy for teaching science and CT, r_{rm} = .551, p = .002. The repeated-measures correlation result is depicted in Figure 3.

Figure 3

The Correlation between Self-Efficacy for Teaching Science and CT



RQ3: Sources of Teachers' Self-Efficacy for Teaching Science in a Computationally Rich Environment

Based on thematic analysis, we found that three sources contributed to teachers' self-efficacy for teaching science in a computationally rich environment: (a) computer programming

experience, (b) students' interest and responses during activities, and (c) teaching repetition and field experience. We elaborate on these three sources below, also showing how teachers' self-efficacy for teaching science developed based on these three sources from pretraining to postteaching.

Computer Programming Experience

We found that most participating teachers worried about teaching science in a computationally rich environment because they were unfamiliar with computer programming. Vivian clearly stated, "For now, I am not confident teaching science using computer programming because I really don't have any experience learning coding itself" (pretraining). Some teachers, particularly Wilma and Regina, knew the term *computational thinking* from a PD in which they had participated. However, prior to this study, they never understood or implemented it in practice. They also were unsure how to teach science in a CS-CT-infused learning environment. Wilma expressed,

> I have heard of computational thinking from a PD in which I participated. However, I never experienced it or saw what CT-integrated science learning looks like in practice. I just know that kind of teaching practice exists, but I have never really tried it out in my classroom. (pretraining interview)

Shirley barely felt confident with integrating computer programming into her science teaching and reflected on her previous experience implementing STEM-integrated activities. She stated, "I had a challenging time when the school asked us to implement STEM-integrated instructional practices because it is just hard to combine mathematics, science, and arts together at once" (pretraining). While most of the teachers did not feel confident teaching science with computer programming because they never learned how to use it, they were enthusiastic about

learning it. The reason for their interest stemmed from their understanding of the importance of integrating technology into science instruction to help students learn scientific concepts optimally. Some teachers felt that their prior experience with other computer-based tools would carry over to computational modeling. Regina, for example, indicated that she felt confident because she had a successful learning experience from participating in a PD, using the computer simulations from PhET to teach electrical circuits. When she tried the method in her classroom, she received positive responses from students and observed a greater improvement in their conceptual understanding than when she used traditional teaching practices.

After the training, the teachers indicated that they felt more confident using the target methods because they had learned enough computer programming and understood what it looked like in the practice. Some of the teachers also stipulated that they even understood why coding is appropriate for representing scientific concepts, especially with visualizing the abstract scientific concepts through the computational model. Alisha expressed,

> I feel more confident teaching science using coding because it will enable students to understand the concept [food webs]…because it can help students visualize the concepts so they can grasp the concepts better…Also, the first time I heard of coding, I felt worried and thought it would be hard, but, after I learned how to do it, I feel more confident. (posttraining written reflection)

When asked about a specific aspect of the program that increased his self-efficacy, Damian mentioned that learning how to code was new to him. After the training, he understood the logic of programming, making him more confident about teaching science using programming. Damian said, “I believe that I can teach science to students using programming …because even though it [coding] was a new thing to me, I now know how to do it”

(posttraining written reflection). In contrast, Aiden, the only CS teacher in this study, expressed that he felt more confident with the coding part but not with the food web concepts. He reasoned, "…because that is not the area I was prepared for" (preteaching written reflection).

After teaching, teachers largely sustained their higher level of self-efficacy when they confirmed that the programming basics was adequate for the activity they taught. Dina expressed,

> I think as long as we know all the basics [about computer programming], I would feel comfortable and confident teaching science using programming…because that means I am ready to also correctly answer students' questions. (postteaching interview)

Dina's response indicates that teachers in this study felt confident only for this particular activity, though. Although understanding an activity's computer programming components increased their confidence, teachers expressed their self-efficacy for teaching science, and CT depends on students' capabilities and responsiveness during the activities. Furthermore, the teachers still harbored doubts about their capabilities to teach science using computer programming and would likely require repeated practice before feeling prepared to better teach students. The following sources showed how teachers' self-efficacy changed due to students' motivation and abilities, as well as teachers' preparation and practice attempts.

Students' Interests and Responses

Although teachers believed that computationally rich science activities are beneficial to both teachers and students, in the pretraining and posttraining interviews, most teachers doubted students would react positively. Their reasons included that such learning activities might be new to students, and that students would feel awkward and even struggle to learn science in such an environment. They feared these conditions might result in students losing motivation. After the

training, Alisha expressed, “I feel worried because my students are diverse in terms of how they grasp and respond to these new activities, and I still feel confused after this [training]” (posttraining interview). Regina echoed this sentiment in her posttraining written reflection, stating, “It [computationally rich environment] would decrease students’ enthusiasm, especially in seventh grade because it is a new thing to them.” Although Alisha and Regina doubted that students would respond positively, they were confident that these activities would be of interest to students with prior experience in computing and robotics and those aspiring computational-related jobs. For example, Regina expressed the activities were “very interesting and challenging and [could] provide a pathway for students who are interested in computer, computing … and information technology” (posttraining written reflection).

The teachers who predicted less-positive responses from their students during the activities sometimes did not receive the predicted reaction during the classroom activities. In particular, Alisha was surprised when her students reacted more positively than she had expected. She indicated how these unexpected reactions from her students positively impacted her self-efficacy during teaching. Alisha stated, “As I told you [researcher], I felt so nervous about how students would react to this new learning approach…but unexpectedly, the first responses from them were positive, and they showed how enthusiastic they were.” (postteaching interview). However, Alisha did not maintain her confidence throughout the four activities. At one point, she stipulated how students’ silence made her confidence drop, and she wondered, “Did I do something wrong?” (postteaching interview). Shirley resonated with Alisha’s experiences, indicating that she found some students gradually enjoying learning science using computer programming, which positively contributed to her self-efficacy when teaching. Shirley

also found that some students did not understand every step in the activities, although she believed she clearly explained them.

Avery reacted differently to students' responses than Shirley and Alisha. He reported that his students kept asking questions about the computer programming platform. Although Avery believed that students' questions are desirable in terms of their engagement, he also felt anxious because of limited time. Avery stated, "I kind of like ran out of time because many students asked too many questions. I felt this was a challenging situation because I only had minimal time for one activity" (postteaching interview). Other preservice teachers had the same challenges as Avery regarding students' questions and limited time. Tamara said, "On Day [Activity] 3, students felt confused and kept asking questions" (postteaching interview). She even heard students comment, "Why is this so difficult? Why don't we just do what we always do?" (postteaching interview), which was also echoed by Vivian's students. The preservice teachers then indicated how their students' questions influenced their teaching confidence during the activities and impacted their self-efficacy.

Besides exposure to computer programming and students' interests and responses throughout the activities, the teachers also indicated that practice, repetition, and teaching directly elevated their self-efficacy for teaching science in a computationally rich environment. The following source shows such evidence.

Repetition and Field Experience

Teachers generally believed that they were confident with their understanding of scientific concepts, including the food web concepts taught in this unit, because they taught these concepts already or learned about them in their teacher education program. Teachers also believed that having taught the concepts several times improved their confidence. Dina stated,

> Based on everyday experience, I feel confident teaching biological concepts because I have taught the concepts several times. In a week, I could probably teach the same concepts five times. Through this repetition, I always recognize some mistakes in my teaching practice and fix them in the next class I am about to teach. But, this does not apply to teaching computational thinking because I do not have such teaching experience." (pretraining)

As already noted, most teachers increased their confidence after the training, and as they became familiar with block-based programming. These teachers still indicated that they needed to continue practicing before teaching students. After repeating the methods several times, the teachers believed they would be able to familiarize themselves with the environment and have a greater understanding of the connection between computer codes and scientific concepts. Wilma, for example, stipulated that she would "practice several times so that [she] could master how to operate the system [computer programming environment]" (posttraining written reflection). In addition to practice, preservice teachers stated that their experience in observing the in-service teachers teaching science in a computationally rich environment helped increase their teaching confidence. Avery expressed, "I feel more confident because I have observed how the other teacher taught and some mistakes the teacher made, so I feel more comfortable teaching science using coding" (preteaching interview).

Notably, none of the teachers practiced much between the after-training and preteaching of the first activity. They indicated that they started to practice more intently when they were about to teach, which was about one or two days before entering each phase in the Use-Modify-Create progression.

Repetition and the number of classes the teachers taught positively influenced teachers' teaching efficacy. Many teachers agreed that they leveraged the first class they taught as an experimental class. The teachers experimented with several different approaches in this first period to identify various teaching approaches and mistakes, such as those related to time. Many teachers also stipulated that the first class they taught was, in effect, a practice class. For instance, Wilma, who led four classes, indicated how careful and less confident she was when teaching the first class. She gradually increased her confidence with more fluent teaching practices in the other classes. Wilma explained,

> I feel everything became easier after I taught. So, I was so careful in the first period because I did not want to miss anything. Next period, I felt more relaxed, and indeed, I felt my teaching skills improved. The more I taught, the easier and more organized my teaching was. (postteaching interview)

Avery shared the same feelings as Wilma. He said that he encountered some problems and mistakes in the first class. However, he could prevent such a problem from occurring in the following classes because he knew how to address the particular situation. Therefore, Avery and many other teachers felt more confident than they did during the first class.

The experience of science teaching in a computationally rich environment sparked teachers' interest in using such a learning environment for other concepts. The teachers saw the merit of using computational modeling to teach science. They believed that such a learning environment fosters students' creativity, independence, critical thinking, and systems thinking. Shirley expressed,

> From the first activity to the fourth activity, students learned some degree of creativity, literacy, and critical and systems thinking. They learned how to be independent. You

> know, many students nowadays rely too much on the Internet and always ask the Internet about any tasks to which they are assigned. But, this activity really helped them to be more independent and creative because they could not find the answers on the Internet. (postteaching interview)

Moreover, although repetition helped improve teachers' interests and self-efficacy, the benefits seemed to be only applied to the specific teaching and learning activities, which was computational modeling of a food web. The teachers indicated that they might not be as confident if given opportunities to teach the other concepts. Furthermore, the teachers did not feel confident developing a completely new unit due to their unfamiliarity with developing new activities using a computer programming environment.

Discussion and Implications

The current study's primary objectives were to explore the developmental trends and sources of middle-school teachers' self-efficacy for teaching both science and CT. In addressing the first and second research questions, we found a pattern of increasing teachers' self-efficacy for teaching both science and CT throughout the study: from pretraining to postteaching. The most apparent changes in science and CT self-efficacy pertained to pretraining to posttraining and from preteaching to postteaching. We also found that significant changes in science teaching self-efficacy were significantly associated with changes in CT teaching self-efficacy. This result implies that both science and CT self-efficacy has an essential relationship to teaching computationally rich science activities. Additionally, we observed that such changes were due to three sources: computer programming experience, students' characteristics, and teaching repetition and experience.

Our findings have corroborated Bandura's (1997) self-efficacy theory. Specifically, the current study showed how self-efficacy emanates from past or mastery experience and vicarious experience. In addition, their psychological and emotional states, influenced by students' interests and comments during the activities, also shaped fluctuations in self-efficacy. Teachers gained mastery and vicarious experiences primarily from the workshop, given that teachers experienced learning how to operate a block-based computer programming platform (mastery) and saw how it could be integrated into science teaching (vicarious). Such experiences might have elevated both teachers' science and CT teaching efficacy beliefs primarily because it was their first encounter with block-based programming and CT-integrated science instruction. This finding resonates with Ketelhut et al. (2020) and Langbeheim et al. (2020), who also observed an increase in teachers' computer programming and science self-efficacy and their attitudes toward computational modeling activity after a workshop or PD. Our findings contribute to this line of research by measuring teachers' self-efficacy for teaching CT, which is not limited to computer programming but includes ways to leverage such skills to solve a problem (Shute et al., 2017). Furthermore, our findings support the previous studies that call for the development of workshops, training, and PD programs to introduce and aid teachers in learning basic computer programming and integrating it into their science teaching and learning activities.

Vicarious experiences influenced teachers' self-efficacy for teaching science and CT after the workshop and positively impacted their self-efficacy after teaching. This result was expected, as previous studies have shown a significant change in teachers' self-efficacy after they experience teaching directly (d'Alessio, 2018; Rupp & Becker, 2021; Tschannen-Moran & Hoy, 2007). However, we found that this vicarious experience did not necessarily immediately impact the teachers. They needed to teach several times to finally reach their "victory" teaching, which

resulted from experimenting in the previous classes they taught. This finding confirms the substantial role of preparation and repetition in shaping teachers' self-efficacy for teaching in a computationally rich science environment. However, this finding also implies that the training was not impactful enough to bolster teachers' self-efficacy, probably due to the online delivery mode in a short period. These conditions might have forced teachers into experimenting with their first classes. We acknowledge this limitation in the current study. Thus, we suggest future studies to implement a faded scaffolding approach to PD to bolster teachers' self-efficacy and prevent them from experimenting with their first classes (Boulden et al., 2018; Lee et al., 2017; Tofel-Grehl et al., 2018).

Additionally, we found that vicarious experience during teaching is associated with teachers' psychological states and affected by students' interests and responses. In our study, teachers' vicarious experiences not only enabled the successful delivery of materials repeatedly but also the way they handled students' lack of interest and responses. From our interviews with the teachers, we discovered that in-service teachers mostly could solve a problem related to students' interests and challenging questions. In contrast, pre-service teachers had difficulty untangling students' confusion. This finding indicates that the teaching experience of pre-service teachers was insufficient to tackle confidence issues mainly related to managing students' confusion and questions. The gap between pre-service and in-service teachers' management skills, both classroom and self, may also correspond to their mastery and vicarious experiences, given that in-service teachers tend to have more experiences than pre-service teachers. Rupp and Becker (2021) also found that pre-service teachers' mastery experiences contributed to their teaching self-efficacy fluctuations. The implication of this finding suggests that when pre-service teachers are involved in teaching a computationally rich science activity, they should be provided

with an opportunity to observe in-service teachers. Such an option may enable the pre-service teachers to obtain a baseline for the typical number of questions and level of confusion students may face. Furthermore, pre-service teachers can see how to structure the lesson's timing and language. In addition, it is important to note that the rapid shift to online teaching and learning because of the pandemic created another layer of complexity and reduced flow of critical social communication during instruction that may have exacerbated the situation for novice teachers.

We believe that our findings provide theoretical and practical implications for understanding teachers' teaching self-efficacy in computationally rich science activities. Like other integrated learning and teaching activities (e.g., Harrell, 2010; Wei, 2020), self-efficacy for teaching science does not suffice to drive teachers' confidence for teaching computationally intensive science. Furthermore, CT (including computer programming) may have been seen as "out-of-field." Hobbs (2013) explained this view as a boundary-crossing event—something for which the teachers see no obligation to associate. However, we believe teachers should be able to tackle this boundary and reshape their teaching identities as more interdisciplinary. Wei (2020) argued that such active identity reshaping could happen when teachers are introduced to and experience using and teaching science in an integrated environment. In the context of the current study, this would be a computationally rich environment. Therefore, researchers and policymakers are encouraged to develop additional PD programs targeting increasing CT and computer programming exposure and its integration into science instruction. Policymakers should also advocate for schools to support teachers teaching science in computationally rich environments. The support includes providing more funding to participate in PD-focused CT-science integrated instruction and ensuring computing-machine accessibility to teachers and

students (Boulden, 2020). Such support may enrich teachers' exposure to CT and computer programming, as well as teaching experiences, in such an environment.

Conclusion, Limitations, and Future Directions

The current study provides insights into the developmental trends and sources of middle-school teachers' self-efficacy for teaching science and CT in a computationally rich environment. Overall, the results showed an increasing trend of teachers' self-efficacy for teaching science and CT throughout the study. We found that self-efficacy for teaching science and CT were significantly correlated. The main three sources of this increasing trend were (a) exposure to computer programming; (b) students' characteristics, including their interests and responses throughout the activities; and (c) teaching repetition and experiences.

The findings of the current study should be carefully interpreted, given that some limitations exist. Some of these constraints can also be addressed in future studies. The first limitation is related to the small sample size. We acknowledge that the small sample size issue prevented us from performing more rigorous statistical analyses, such as multilevel modeling, to examine the dynamics of teachers' self-efficacy further. Thus, future studies are encouraged to include more teachers (> 50; McNeish & Stapleton, 2017) to enable running additional advanced statistical techniques and exploring the impacts of teachers' characteristics more deeply (e.g., pre-or in-service, years of teaching, computer programming experience) on teachers' self-efficacy development for teaching science and CT in a computationally rich environment. Moreover, the limited sample size in the current study might have impacted the posthoc results, given that we did not find any significant *p* values after Bonferroni corrections. Increasing the sample size may also address this issue. Last, increasing the sample size would allow for gathering richer data on the variety of sources that encourage teachers' science and CT teaching

self-efficacy. Expanding this research can lead to a more thorough understanding of teachers' self-efficacy for teaching in a computationally rich environment—including how teachers' self-efficacy in science and CT influence students' learning science in such an environment.

A second limitation pertains to online training and teaching. These conditions might have interfered with the results, given that most of the teachers were new to a computationally rich science environment and might have been uncomfortable teaching a new topic in a synchronous online environment. The remote instruction prevented the two teachers from moving forward teaching in their classrooms due to their students' limited access to computers and the Internet. Therefore, future studies can try to deliver the workshop in-person and ask teachers to teach in a face-to-face setting. A face-to-face setting may also enable researchers to gather richer data, especially during activity interviews and classroom interactions. Such data are integral in explaining teachers' fluctuating self-efficacy (Rupp & Becker, 2021) and describing in greater detail how teachers shift their self-efficacy across the UMC progression and through multiple rounds of instruction. Furthermore, in-person teaching and learning activities may enable teachers to include all students in CT-integrated science activities, given that computers and the Internet are available to them in schools.

Future studies are also encouraged to implement more extended training or PD programs and diversify the computer programming environments. These approaches may increase teachers' exposure to computer programming and elevate their self-efficacy through mastery and vicarious experiences (Rich et al., 2021). Additionally, future studies may develop an instrument to measure teachers' self-efficacy for teaching in a computationally rich science environment. With such a measure, researchers can run predictive statistical models to examine which type of

teaching self-efficacy—science or CT—affects the variability more significantly in combined science–CT teaching self-efficacy.

Acknowledgement

We thank the eleven teachers participating in this study. We also thank Dr. Jessica Vandenberg for her constructive feedback on the initial draft of this work, and Ai Nurlaelasari Rusmana for being the second coder. This material is based upon the work supported by the National Science Foundation under Grant No. DRL-1640141. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The portions of this study were also supported by North Carolina State University's College of Education and the Friday Institute for Educational Innovation through Dissertation Research Continuity Award and Graduate Student Fellow Award, respectively.

REFERENCES

Adler, R. F., & Kim, H. (2018). Enhancing future K-8 teachers' computational thinking skills through modeling and simulations. *Education and Information Technologies*, *23*(4), 1501–1514. https://doi.org/10.1007/s10639-017-9675-1

Agresti, A. (2018). *Statistical methods for the social sciences (global edition)*. London, UK: Pearson Education.

Aksit, O., & Wiebe, E. N. (2020). Exploring force and motion concepts in middle grades using computational modeling: A classroom intervention study. *Journal of Science Education and Technology*, *29*(1), 65–82. https://doi.org/10.1007/s10956-019-09800-z

Alt, D. (2018). Science teachers' conceptions of teaching and learning, ICT efficacy, ICT professional development and ICT practices enacted in their classrooms. *Teaching and Teacher Education*, *73*, 141–150. https://doi.org/10.1016/j.tate.2018.03.020

Althauser, K. (2015). Job-embedded professional development: Its impact on teacher self-efficacy and student performance. *Teacher Development*, *19*(2), 210–225. https://doi.org/10.1080/13664530.2015.1011346

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, *8*, 1–13. https://doi.org/10.3389/fpsyg.2017.00456

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman and Company.

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, *1*(2), 164–180. https://doi.org/10.1111/j.1745-6916.2006.00011.x

Bandura, A., (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

Betebenner, D. (2009). Norm-and criterion-referenced student growth. *Educational Measurement: Issues and Practice*, *28*(4), 42–51. https://doi.org/10.1111/j.1745-3992.2009.00161.x

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really?. *Educational Psychology Review*, *15*(1), 1–40. https://doi.org/10.1023/A:1021302408382

Boulden, D. C. (2020). *Building the capacity of in-service teachers to integrate and teach computational thinking* [Doctoral dissertation, North Carolina State University]. NC State Libraries.

Boulden, D. C., Rachmatullah, A., Oliver, K. M., & Wiebe, E. (2021). Measuring in-service teacher self-efficacy for teaching computational thinking: development and validation of the T-STEM CT. *Education and Information Technologies*, 1–27. https://doi.org/10.1007/s10639-021-10487-2

Boulden, D. C., Wiebe, E., Akram, B., Aksit, O., Buffum, P. S., Mott, B., Boyer, K. E. & Lester, J. (2018). Computational thinking integration into middle grades science classrooms: Strategies for meeting the challenges. *Middle Grades Review*, *4*(3).

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNalley & Co.

Campbell, T., Longhurst, M. L., Wang, S. K., Hsu, H. Y., & Coster, D. C. (2015). Technologies and Reformed-Based Science Instruction: The Examination of a Professional Development Model Focused on Supporting Science Teaching and Learning with Technologies. *Journal*

of Science Education and Technology, *24*(5), 562–579. https://doi.org/10.1007/s10956-015-9548-6

Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: issues and techniques. *Journal of Advanced Nursing*, *58*(4), 386–395. https://doi.org/10.1111/j.1365-2648.2007.04242.x

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, *18*(8), 947–967. https://doi.org/10.1016/S0742-051X(02)00053-7

Creswell, J. W., & Clark, V. L. P. (2018). *Designing and conducting mixed methods research: Third edition*. Thousand Oaks, CA: Sage Publications.

d'Alessio, M. A. (2018). The Effect of Microteaching on Science Teaching Self-Efficacy Beliefs in Preservice Elementary Teachers. *Journal of Science Teacher Education*, *29*(6), 441–467. https://doi.org/10.1080/1046560X.2018.1456883

Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C. D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education*, *24*(3), 751–766. https://doi.org/10.1016/j.tate.2007.02.010

Gallup Inc. & Google Inc. (2016). *Diversity gaps in computer science: exploring the underrepresentation of girls, Blacks and Hispanics*. Retrieved from http://goo.gl/PG34aH.

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, *76*(4), 569–582. https://doi.org/10.1037/0022-0663.76.4.569

Grover, S., & Pea, R. (2018). Computational thinking: A competency whose time has come. In S. Sentence, E. Barendsen., & C. Schulte (Eds.) *Computer Science Education:*

Perspectives on Teaching and Learning in School (pp. 19–38). London, UK: Bloomsbury Academic.

Hadjiachilleos, S., Avraamidou, L., & Papastavrou, S. (2013). The Use of Lego Technologies in Elementary Teacher Preparation. *Journal of Science Education and Technology*, *22*(5), 614–629. https://doi.org/10.1007/s10956-012-9418-4

Harrell, P. E. (2010). Teaching an integrated science curriculum: Linking teacher knowledge and teaching assignments. *Issues in Teacher Education*, *19*(1), 145–165.

Ho, A. D. (2008). The problem with "proficiency": Limitations of statistics and policy under No Child Left Behind. *Educational Researcher*, *37*(6), 351–360. https://doi.org/10.3102/0013189X08323842

Hobbs, L. (2013). Teaching 'out-of-field' as a boundary-crossing event: Factors shaping teacher identity. *International Journal of Science and Mathematics Education*, *11*(2), 271–297. https://doi.org/10.1007/s10763-012-9333-4

Irgens, G. A, Dabholkar, S., Bain, C., Woods, P., Hall, K., Swanson, H., Horn, M., & Wilensky, U. (2020). Modeling and measuring high school students' computational thinking practices in science. *Journal of Science Education and Technology*, *29*(1), 137–161. https://doi.org/10.1007/s10956-020-09811-1

Jaipal-Jamani, K., & Angeli, C. (2017). Effect of Robotics on Elementary Preservice Teachers' Self-Efficacy, Science Learning, and Computational Thinking. *Journal of Science Education and Technology*, *26*(2), 175–192. https://doi.org/10.1007/s10956-016-9663-z

Jong, C., & Hodges, T. E. (2013). The influence of elementary pre-service teachers' mathematical experiences on their attitudes towards teaching and learning

mathematics'. *International Electronic Journal of Mathematics Education*, *8*(2-3), 100–122.

Ketelhut, D. J., Mills, K., Hestness, E., Cabrera, L., Plane, J., & McGinnis, J. R. (2020). Teacher change following a professional development experience in integrating computational thinking into elementary science. *Journal of Science Education and Technology*, *29*(1), 174–188. https://doi.org/10.1007/s10956-019-09798-4

Lachlan-Haché, M., & Castro, M. (2015). *Proficiency or growth: An exploration of two approaches for writing student learning targets*. American Institutes for Research, Retrieved from http://www.air.org/sites/default/files/Explorationof-Two-Approaches-Student-Learning-Targets-April-2015.pdf.

Lakshmanan, A., Heath, B. P., Perlmutter, A., & Elder, M. (2011). The impact of science content and professional learning communities on science teaching efficacy and standards-based instruction. *Journal of Research in Science Teaching*, *48*(5), 534–551. https://doi.org/10.1002/tea.20404

Langbeheim, E., Perl, D., & Yerushalmi, E. (2020). Science Teachers' Attitudes towards Computational Modeling in the Context of an Inquiry-Based Learning Module. *Journal of Science Education and Technology*, *29*(6), 785–796. https://doi.org/10.1007/s10956-020-09855-3

Lauermann, F., & Karabenick, S. A. (2011). Taking teacher responsibility into account (ability): Explicating its multiple components and theoretical status. *Educational Psychologist*, *46*(2), 122–140. https://doi.org/10.1080/00461520.2011.558818

Lee, H., Longhurst, M., & Campbell, T. (2017). Teacher learning in technology professional development and its impact on student achievement in science. *International Journal of Science Education*, *39*(10), 1282–1303. https://doi.org/10.1080/09500693.2017.1327733

Lee, I., & Malyn-Smith, J. (2020). Computational thinking integration patterns along the framework defining computational thinking from a disciplinary perspective. *Journal of Science Education and Technology*, *29*(1), 9–18. https://doi.org/10.1007/s10956-019-09802-x

Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., Malyn-Smith, J., & Werner, L. (2011). Computational thinking for youth in practice. *ACM Inroads*, *2*(1), 32–37. https://doi.org/10.1145/1929887.1929902

Leonard, J., Buss, A., Gamboa, R., Mitchell, M., Fashola, O. S., Hubert, T., & Almughyirah, S. (2016). Using robotics and game design to enhance children's self-efficacy, stem attitudes, and computational thinking skills. *Journal of Science Education and Technology*, *25*(6), 860–876. https://doi.org/10.1007/s10956-016-9628-2

Lytle, N., Cateté, V., Boulden, D., Dong, Y., Houchins, J., Milliken, A., ... & Barnes, T. (2019, July). Use, Modify, Create: Comparing computational thinking lesson progressions for STEM classes. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education* (pp. 395–401).

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). Evaluating value-added models for teacher accountability. *PsycEXTRA Dataset*. https://doi.org/10.1037/e658712010-001.

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review*, *28*(2), 295–314. https://doi.org/10.1007/s10648-014-9287-x

Menon, D., & Sadler, T. D. (2016). Preservice Elementary Teachers' Science Self-Efficacy Beliefs and Science Content Knowledge. *Journal of Science Teacher Education*, *27*(6), 649–673. https://doi.org/10.1007/s10972-016-9479-y

Menon, D., Chandrasekhar, M., Kosztin, D., & Steinhoff, D. C. (2020). Impact of mobile technology-based physics curriculum on preservice elementary teachers' technology self-efficacy. *Science Education*, *104*(2), 252–289. https://doi.org/10.1002/sce.21554

National Research Council (NRC). (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.

NGSS Lead States. (2013). *Next generation science standards. For states, by states*. Washington, DC: The National Academies Press.

Pajares (2002). *Self-efficacy beliefs in academic contexts: An outline*. Retrieved March, 25, 2021, from https://www.uky.edu/~eushe2/Pajares/efftalk.html.

Peel, A., Sadler, T. D., & Friedrichsen, P. (2019). Learning natural selection through computational thinking: Unplugged design of algorithmic explanations. *Journal of Research in Science Teaching*, *56*(7), 983–1007. https://doi.org/10.1002/tea.21545

Rachmatullah, A., Mayhorn, C. B., & Wiebe, E. N. (2021). The effects of prior experience and gender on middle school students' computer science learning and monitoring accuracy in the Use-Modify-Create progression. *Learning and Individual Differences*, *86*, 101983. https://doi.org/10.1016/j.lindif.2021.101983

Rich, P. J., Mason, S. L., & O'Leary, J. (2021). Measuring the effect of continuous professional development on elementary teachers' self-efficacy to teach coding and computational thinking. *Computers & Education*, 104196. https://doi.org/10.1016/j.compedu.2021.104196

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, *74*(6), 625–637. https://doi.org/10.1002/sce.3730740605

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*(1), 1–28. https://doi.org/10.1037/h0092976

Rupp, D., & Becker, E. S. (2021). Situational fluctuations in student teachers' self-efficacy and its relation to perceived teaching experiences and cooperating teachers' discourse elements during the teaching practicum. *Teaching and Teacher Education*, *99*, 103252. https://doi.org/10.1016/j.tate.2020.103252

Saldaña, J. (2013). *The coding manual for qualitative researchers (2nd ed.)*. Thousand Oaks, CA: SAGE Publications Inc.

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

Srisuradetchai, P., & Trakultraipruk, N. (2016). Skillings-Mack Statistic: Computer-Intensive Methods. *Journal of Applied Statistics and Information Technology*, *1*(2), 33–45.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research techniques*. Thousand Oaks, CA: Sage publications.

Swanson, H., Anton, G., Bain, C., Horn, M., & Wilensky, U. (2019). Introducing and assessing computational thinking in the secondary science classroom. In S-C. Kong & H. Abelson, *Computational Thinking Education* (pp. 99–117). Singapore: Springer.

Tang, X., Yin, Y., Lin, Q., Hadad, R., & Zhai, X. (2020). Assessing computational thinking: A systematic review of empirical studies. *Computers & Education*, *148*(January), 1–22. https://doi.org/10.1016/j.compedu.2019.103798

Thomson, M. M., Huggins, E., & Williams, W. (2019). Developmental science efficacy trajectories of novice teachers from a STEM-Focused program: A longitudinal mixed-methods investigation. *Teaching and Teacher Education*, *77*, 253–265. https://doi.org/10.1016/j.tate.2018.10.010

Tofel-Grehl, C., Searle, K. A., & Feldon, D. F. (2018). Professional development for secondary science teachers: A faded scaffolding approach to preparing teachers to integrate computing. In *Proceedings of International Society of the Learning Sciences*. Retrieved from https://repository.isls.org/bitstream/1/903/1/69.pdf

Tschannen-Moran, M., & Barr, M. (2004). Fostering student learning: The relationship of collective teacher efficacy and student achievement. *Leadership and Policy in Schools*, *3*(3), 189–209. https://doi.org/10.1080/15700760490503706

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. https://doi.org/10.1016/S0742-051X(01)00036-1

Tschannen-Moran, M., & McMaster, P. (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. *The Elementary School Journal*, *110*(2), 228–245.

Unfried, A., Rachmatullah, A., Alexander, A., & Wiebe, E. (under review). T-STEM Science Scale: Alternative to STEBI.

Wagh, A., Cook-Whitt, K., & Wilensky, U. (2017). Bridging inquiry-based science and constructionism: Exploring the alignment between students tinkering with code of computational models and goals of inquiry. *Journal of Research in Science Teaching*, *54*(5), 615–641. https://doi.org/10.1002/tea.21379

Wei, B. (2020). An exploratory study of teacher development in the implementation of integrated science curriculum. *Research in Science Education*, *50*(6), 2189–2206. https://doi.org/10.1007/s11165-018-9768-x

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

Williams, D. M., Anderson, E. S., & Winett, R. A. (2005). A review of the outcome expectancy construct in physical activity research. *Annals of Behavioral Medicine*, *29*(1), 70–79. https://doi.org/10.1207/s15324796abm2901_10

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

Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *366*(1881), 3717–3725. https://doi.org/10.1098/rsta.2008.0118

Zangori, L., Foulk, J., Sadler, T. D., & Peel, A. (2018). Exploring elementary teachers' perceptions and characterizations of model-oriented issue-based teaching. *Journal of*

Science Teacher Education, *29*(7), 555–577. https://doi.org/10.1080/1046560X.2018.1482173

Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015. https://doi.org/10.3102/0034654315626801

CHAPTER 5: Conclusion

This chapter provides a summary and contributions of each study discussed in the previous three chapters. This chapter also offers suggestions for future research based on the findings and limitations emerging from these studies.

Summary

Chapter 2

This study was motivated by previous studies hypothesizing that the emerging instructional approaches where students are engaged in building a computational model of a scientific phenomenon could increase their conceptual understanding around that scientific phenomenon, including ecosystems (Nguyen & Santagata, 2020). In addition to developing a conceptual understanding of the food webs, some studies also found that engaging students in a computationally rich environment had been found to advance students' systems thinking skills (Mahaffy et al., 2019; Nguyen & Santagata, 2020). Chapter 2 addresses this dissertation's first research question around the impacts of computational modeling activity on middle school students' computational thinking, systems thinking, and conceptual understanding of food webs.

The study in Chapter 2 employed a quasi-experimental design where students were divided into either a computational modeling condition or a paper-based pictorial modeling condition. They learned the same food web concepts within the same amount of time which is 4 x 60 minutes. All the activities were conducted online. Students in both conditions took the same assessments before and after the activities. They took the CTA-M assessment (Wiebe et al., 2019) and systems thinking-embedded food web assessment (Mambrey et al., 2020). In addition, they took a formative assessment after each activity in which they answered an open-ended question asking what they thought they learned from each activity. Multilevel modeling and

repeated-measures correlation tests were used to analyze the quantitative data. Epistemic Network Analysis (ENA; Shaffer & Ruis, 2017) was performed to analyze students' responses to the open-ended question.

The results revealed that there was a significant increase in students' systems thinking skills. A closer analysis was performed on each of the conceptualized systems thinking components. A significant increase was found in two out of three systems thinking components: systems organization and systems behavior. These significant increases happened to all students in both conditions. The significance disappeared when students' SES and academic scores were added to the analyses. Regarding CT skills, the results showed a significant increase from pre- to post-test only for students in the computational modeling condition. This result remained significant even after controlling for SES and academic scores. Results from correlation tests indicated that CT and systems thinking were weakly correlated, diverging from prior hypothesized relationships (Berland & Wilensky, 2015; Weintrop et al., 2016). ENA results also showed no significant difference in students' statements regarding systems thinking and science concepts between the two conditions. In contrast, a significant difference was found in the frequency of students' expression of CS/CT concepts. That is, the results indicated that students in the computational modeling condition said they learned both the food web and CS/CT concepts.

Chapter 3

The study presented in Chapter 3 was driven by the increasing and growing number of computationally intensive science careers – jobs that require expertise in both science content knowledge and the ability to use CS tools and methods (National Academies of Sciences, Engineering, and Medicine, 2018). A thorough literature review indicated a lack of an instrument

to measure K-12 students' interests in such careers. Thus, this study aimed to develop and validate such an instrument, called the CISCI (Computationally Intensive Science Career Interests) instrument. This instrument was intended for middle school students, given that this age group is at a critical juncture when they start to crystalize and identify interests in certain careers (Skamp, 2007; Wiebe et al., 2018). In addition, this study explored the significant predictors of students' interests in computationally intensive science (CIS) careers and examined the impact of participating in a computationally rich science activity on students' career interests.

Lent et al.'s (2002) social cognitive career theory (SCCT) of interest model was used as a theoretical framework in this study. The SCCT interest model emphasizes the importance of cognitive and experiential factors that influence individuals' interests in certain careers. This model also describes the interrelationship between interests and motivation to acquire certain skills and behavior related to such careers. Guided by the SCCT of interest model, the CISCI instrument was hypothesized to consist of five constructs (construct, factor, and dimension are used interchangeably): CIS career value, CIS academic self-efficacy, CIS career self-efficacy, CIS career interest and goal, and CIS parental supports and role models. The development and validation process of the CISCI instrument was guided by AERA et al.'s (2014) *Standards for Educational and Psychological Testing*. This study focused on four aspects of validity: test content, consequences of testing, internal structure, and criterion validity.

A total of fourteen expert panelists were involved in assessing the initial CISCI instrument, quantitatively and qualitatively. For content validity, Lawshe's (1975) content validity ratio (CVR) was used to determine each item's essentiality based on the experts' feedback. Seventeen items were retained after the content validity process and were administered to 934 middle school students. Exploratory Factor Analysis (EFA) and Multidimensional Rasch

modeling were performed to assess the dimensionality and the quality of the instrument. EFA suggested two- and three-factor solutions. The EFA results were then compared with one-factor (baseline) and five-factor (hypothesized model) models in Multidimensional Rasch Modeling. The three- and five-dimension models turned out to have fewer misfitting items than the other models. Two items were identified as misfitting and then were removed from the instrument, resulting in a total of 15 items. Confirmatory Factor Analysis (CFA) was used to compare the statistical structure of the two models. The CFA results revealed that the five-factor model was the best fit and was therefore used in the subsequent analyses.

Multiple linear regression tests were used to identify the significant predictors for the five constructs in the CISCI instrument, addressing aspects of criterion validity. The results showed that middle school students' interests in CIS careers were positively associated with their science attitudes, CS attitudes, CT skills, math scores, and prior experience in CS-related activities. The paired-sample *t*-tests were performed to investigate the five constructs' changes before and after 186 students participated in a week-long computationally-rich science activity. The results indicated a significant increase in students' exposure to CIS parental support and role models.

Chapter 4

The study in Chapter 4 was motivated to elicit the developmental trends and sources of middle school teachers' self-efficacy for teaching science and CT in a computationally rich environment. Integrated science teaching and learning has been studied in the science education field based on its value in promoting more holistic instructional activities (Wei, 2020). One specific variation is CT-integrated science instruction, where students and teachers learn and teach science using computer programming (Hadjiachilleos et al., 2013; Langbeheim et al., 2020). While this approach has demonstrated positive influence on students' cognitive, affective-

motivational, and practices, as demonstrated in Chapters 2 and 3, the research literature still lacks understanding of how teachers build their confidence when teaching science in such an environment. Most teachers have no prior experience with instruction that integrates science and CT/CS concepts (Ketelhut et al., 2020). A lack of such knowledge may influence teachers' teaching efficacy. Exploring how teachers develop their self-efficacy for teaching in a CT-integrated science learning environment and identifying the sources of their self-efficacy may be central for forwarding students' learning in this area (Althauser, 2015; Zee & Koomen, 2016).

The study in Chapter 4 employed a mixed-methods approach by collecting quantitative data on teachers' self-efficacy for teaching science and CT and qualitative data in the form of interviews and written reflections. A total of eleven teachers participated in this study. They were engaged in a CT-science integrated workshop for five hours, after which they taught the materials to their students. The materials were about teaching the food web concepts using block-based programming. The teachers took two questionnaires four times: before and after the training/workshop and before and after teaching. They were also interviewed or provided written reflections on their confidence with teaching science in a computationally-rich environment at those four time points. Skillings-Mack and repeated-measures correlation tests were used to examine the changes in teachers' self-efficacy across four time points and investigate the association between science and CT teaching efficacy beliefs. Thematic analysis was used to gather evidence of the sources of teachers' self-efficacy.

The results indicated significant changes in science and CT teaching efficacy beliefs from before the workshop and after teaching. A substantial increase was detected from before workshop to after workshop and from before teaching to after teaching. The results also showed a significant repeated association between self-efficacy for teaching science and CT. Based on

the qualitative thematic analysis, three sources were found that might explain such changes: (1) teachers' experiences with computer programming/coding, (2) students' interests in and responses to the computationally rich science activities, and (3) teachers' experiences in teaching science in a computationally-rich environment several times.

Contributions of Research

The three studies in this dissertation contribute to the literature and practice by advancing how students and teachers learn and teach science in a computationally-rich environment. The first study highlights the affordances of computational modeling activities in forwarding all dimensions of 3D instructional design. Research has found that both computational modeling and paper-based pictorial modeling can improve students' systems thinking skills and conceptual understanding of food webs (e.g., Nguyen & Santagata, 2020). The first study extends prior findings by providing additional evidence for computational modeling being an appropriate science learning activity that also enables students to advance their CT skills. The first study provides an exemplary science learning activity covering the three dimensions in NGSS–disciplinary content knowledge, crosscutting concepts, and science and engineering practices. This exemplary science learning activity contributes to the practice as it can be used by the teachers, especially those new to computer programming. The first study's activities leverage a block-based programming language that is relatively accessible and easy to learn and teach for beginning programmers. The study also provides further clarification of the relationship of CT and systems thinking that has been hypothesized but not extensively empirically documented in the literature (Berland & Wilensky, 2015; Weintrop et al., 2016).

The second study contributes to the literature by providing a validated and psychometrically sound instrument to measure middle school students' interests in

computationally intensive science careers. The second study provides evidence for the importance of students' science and CS attitudes as well as prior CS-related experience to developing students' interests in such careers. The results also suggest the influence of students' CT and mathematics achievement in predicting their career interests. Moreover, the second study shows the impact of participating in a computational modeling activity on students' career interests. This work also uncovered an interesting impact through the increased the frequency of talk between students and parents about CIS-related careers. Taken together, the second study's findings support Lent et al.'s (2002) social cognitive career theory, especially regarding the influence of demographics, cognitive, affective-motivational, and experience factors in predicting students' career interests.

The third study has theoretical implications by supporting Bandura's self-efficacy theory (1997) as it relates to teacher self-efficacy. The third study underscores the importance of professional development, training, and workshop programs as well as their teaching experience to elevating middle school teachers' science and CT teaching self-efficacy. Such experiences are crucial, especially for teachers who are new to teaching science in a computationally rich environment. The study suggests that teachers with no prior experience teaching computational modeling gain confidence from their exposure to block-based computer programming, their students' reactions and interests, and their repeated teaching experiences. Furthermore, the third study contributes to the literature by providing a piece of evidence that teachers need both science and CT teaching self-efficacy when they teach science in a computationally rich environment.

While the pandemic-driven need for teachers to engage in PD and teach online, and for students to rely exclusively on this distance-based instruction, these constraints allowed us to

demonstrate the flexibility and efficacy of computational modeling as an instructional strategy in such conditions. It also pointed to, perhaps serendipitously, the value of vicarious interactions of students and family members during home-based instruction; this perhaps led to both exposure to new scientific methods utilizing computational modeling and enhanced communication between students and parents about learning and future careers in this area. On the other hand, the studies in this dissertation also demonstrated how the pandemic, internationally, has continued to exacerbate the digital divide, as classrooms with students lacking a home computer could not participate in this study.

Limitations and Future Directions

Like many studies, the three studies in this dissertation have some limitations that compel the readers to interpret the findings carefully. Future studies are encouraged to address these limitations and elaborate on these results in ways that contribute to the literature. The first limitation is the use of non-random sampling. All the students who participated in the studies were the result of convenience sampling at the teacher and school level. Even though the proportions of some demographic variables, such as gender and age, in the samples align with the national demographic population profiles, students' SES levels' ratios did not. The students from high SES families were over-represented in the final dataset. This over-representation may also help explain the high percentage of students with above-average academic scores. This phenomenon is common, especially in the online learning environment (Deschacht & Goeman, 2015; Lee et al., 2013). Therefore, future studies are encouraged to replicate the intervention in Chapters 2 and 3, but in an in-person mode where resource factors such as computer ownership will not bias the sample. In a face-to-face setting, government subsidies of computing technology infrastructure may provide universal access and provide opportunities for students of all SES

levels to participate. The in-person setting may also enable researchers to collect richer qualitative data such as during activity interviews, classroom interactions, and situational observations that can be used to substantiate and enrich more distal data collection techniques.

The online learning and teaching environment due to the COVID-19 pandemic might have also shaped the results in Chapter 4. Most of the teachers were new to teaching science in a computationally rich environment, and together with the online instructional environment, they might not be comfortable enough teaching such novel material. Similarly, the online workshop might have prevented teachers from grasping all the materials, leading to unprepared materials when teaching the materials to the students. Another limitation regarding the study in Chapter 4 is that the online environment had made two of the teachers withdraw from participation in the study due to their students' limited access to the Internet and computer. Hence, a face-to-face setting is integral to gather more evidence for the study in Chapter 4. Future studies can also collect data on teachers' fluctuating self-efficacy based on their gestures or behaviors in an in-person environment. These data are integral to provide a more holistic description of teachers' science and CT teaching self-efficacy.

Additionally, more extended and alternative configurations of the interventions are encouraged to be studied in the future. The computational modeling activity used in this dissertation employed the Use-Modify-Create progression instructional strategy in the Snap! programming environment. Thus, the findings are to some degree limited to this specific learning context. Future studies may use different instructional strategies such as the combinations of (plugged) computational modeling and (unplugged) pictorial modeling, or different programming languages and platforms. In addition to the specific components of computational modeling, future studies may also explore modeling with different scientific concepts. This

extension would allow the researchers to both generalize findings of this study and possibly longitudinally explore the development of students' systems thinking, CT, and their interests in computationally intensive careers. Such longitudinal investigation is crucial to identify the critical point in which students have significant inflections in their career interests and CT skills.

An important limitation of the teacher study in Chapter 4 is related to sample size. The small sample size prevented us from running more sophisticated statistical analyses, such as multilevel modeling, and examining quantitatively the inter- and intra-variability in teachers' science and CT teaching self-efficacy. Future studies can address this limitation by increasing the sample size to at least 50 (McNeish & Stapleton, 2017) in order to better explore the fluctuation and differences between teachers' science and CT teaching self-efficacy. Increasing the sample size for teachers in the overall intervention study may also help better examine students' systems thinking and conceptual understanding of food web concepts in Chapter 2's study. By increasing the teachers' sample size, there would be a corresponding increase in the diversity of students in the sample, and teachers' characteristics could possibly be controlled by more advanced statistical techniques. Future studies can also increase the number of schools so that schools' characteristics can be added to the statistical models when examining the efficacy of computational modeling activities and teachers' self-efficacy fluctuations.

Regarding the CT instrument used in Chapters 2 and 3, future studies can compare an instrument that is free from utilizing a programming context for items to the one that uses non-programming representations as a basis for the items. In addition, the items may also utilize other programming languages including text-based ones. Such comparison may enable researchers to investigate the impact of a programming language in measuring individuals' CT skills. In addition, future studies are also encouraged to measure students' CT dispositions (Hava & Ünlü,

2021) to gather evidence for the efficacy of computationally rich science activities on improving students' CT skills and dispositions.

Similarly, future studies can measure students' conceptual understanding of the food web conceptual understanding and systems thinking separately. This would allow researchers to look closely at the relationship between the two constructs and CT. One possible instrument to measure systems thinking separate from the science concept is utilizing an instrument measuring cyclic thinking (Ben-Zvi Assaraf & Orion, 2005) or using different theoretical frameworks such as structure-function-behavior (Hmelo-Silver et al., 2017) or mechanistic reasoning (Russ et al., 2008). Moreover, future studies can use qualitative interview data to corroborate the findings from the quantitative assessments.

Last, regarding teachers' PD or workshop programs, the findings from Chapter 5 suggest that extended programs are needed to increase science teachers' exposure to computer programming, and how it is best supported in the context of computational modeling. Such comprehensive PD programs would also allow the teachers to be better prepared for teaching science in a computationally rich environment. One benefit would be that teachers may have more mastery and vicarious experiences that facilitate the increase in their science and CT teaching self-efficacy (Rich et al., 2021). Moreover, teachers' science and CT teaching self-efficacy may be measured with a single, unified instrument developed to measure science teachers' teaching self-efficacy in a computationally rich environment. Such an instrument would reduce teachers' testing fatigue, and thus teachers can concentrate more on responding to the questionnaire if it is measured repeatedly.

REFERENCES

Aksit, O., & Wiebe, E. N. (2020). Exploring force and motion concepts in middle grades using computational modeling: A classroom intervention study. *Journal of Science Education and Technology*, *29*(1), 65–82. https://doi.org/10.1007/s10956-019-09800-z

Althauser, K. (2015). Job-embedded professional development: Its impact on teacher self-efficacy and student performance. *Teacher Development*, *19*(2), 210–225. https://doi.org/10.1080/13664530.2015.1011346

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman and Company.

Ben-Zvi-Assaraf, O., & Orion, N. (2005). Development of system thinking skills in the context of earth system education. *Journal of Research in Science Teaching*, *42*(5), 518–560. https://doi.org/10.1002/tea.20061

Berland, M., & Wilensky, U. (2015). Comparing virtual and physical robotics environments for supporting complex systems and computational thinking. *Journal of Science Education and Technology*, *24*(5), 628–647. https://doi.org/10.1007/s10956-015-9552-x

Deschacht, N., & Goeman, K. (2015). The effect of blended learning on course persistence and performance of adult learners: A difference-in-differences analysis. *Computers & Education*, *87*, 83–89. https://doi.org/10.1016/j.compedu.2015.03.020

Hadjiachilleos, S., Avraamidou, L., & Papastavrou, S. (2013). The Use of Lego Technologies in Elementary Teacher Preparation. *Journal of Science Education and Technology*, *22*(5), 614–629. https://doi.org/10.1007/s10956-012-9418-4

Hava, K., & Ünlü, Z. K. (2021). Investigation of the relationship between middle school students' computational thinking skills and their STEM career interest and attitudes toward inquiry. *Journal of Science Education and Technology*, 1–12. https://doi.org/10.1007/s10956-020-09892-y

Hmelo-Silver, C. E., Jordan, R., Eberbach, C., & Sinha, S. (2017). Systems learning with a conceptual representation: a quasi-experimental study. *Instructional Science*, *45*(1), 53–72. https://doi.org/10.1007/s11251-016-9392-y

Jin, H., Shin, H. J., Hokayem, H., Qureshi, F., & Jenkins, T. (2019). Secondary students' understanding of ecosystems: A learning progression approach. *International Journal of Science and Mathematics Education*, *17*(2), 217–235. https://doi.org/10.1007/s10763-017-9864-9

Ketelhut, D. J., Mills, K., Hestness, E., Cabrera, L., Plane, J., & McGinnis, J. R. (2020). Teacher change following a professional development experience in integrating computational thinking into elementary science. *Journal of Science Education and Technology*, *29*(1), 174–188. https://doi.org/10.1007/s10956-019-09798-4

Langbeheim, E., Perl, D., & Yerushalmi, E. (2020). Science Teachers' Attitudes towards Computational Modeling in the Context of an Inquiry-Based Learning Module. *Journal of Science Education and Technology*, *29*(6), 785–796. https://doi.org/10.1007/s10956-020-09855-3

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575.

Lee, Y., Choi, J., & Kim, T. (2013). Discriminating factors between completers of and dropouts from online learning courses. *British Journal of Educational Technology*, *44*(2), 328–337. https://doi.org/10.1111/j.1467-8535.2012.01306.x

Lent, R.W., Brown, S. D., & Hackett, G. (2002). Social cognitive career theory. In D. Brown, *Career choice and development fourth edition* (pp. 255–311). San Francisco, CA: Jossey-Bass.

Mahaffy, P. G., Matlin, S. A., Holme, T. A., & MacKellar, J. (2019). Systems thinking for education about the molecular basis of sustainability. *Nature Sustainability*, *2*(5), 362–370. https://doi.org/10.1038/s41893-019-0285-3

Mambrey, S., Timm, J., Landskron, J. J., & Schmiemann, P. (2020). The impact of system specifics on systems thinking. *Journal of Research in Science Teaching*, 1–20. https://doi.org/10.1002/tea.21649

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review*, *28*(2), 295–314. https://doi.org/10.1007/s10648-014-9287-x

Mehren, R., Rempfler, A., Buchholz, J., Hartig, J., & Ulrich-Riedhammer, E. M. (2018). System competence modelling: Theoretical foundation and empirical validation of a model involving natural, social and human-environment systems. *Journal of Research in Science Teaching*, *55*(5), 685–711. https://doi.org/10.1002/tea.21436

National Academies of Sciences, Engineering, and Medicine. (2018). *Assessing and responding to the growth of computer science undergraduate enrollments*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24926.

Nguyen, H., & Santagata, R. (2020). Impact of computer modeling on learning and teaching systems thinking. *Journal of Research in Science Teaching*, 1–28. https://doi.org/10.1002/tea.21674

Rich, P. J., Mason, S. L., & O'Leary, J. (2021). Measuring the effect of continuous professional development on elementary teachers' self-efficacy to teach coding and computational thinking. *Computers & Education*, 104196. https://doi.org/10.1016/j.compedu.2021.104196

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, *92*(3), 499–525. https://doi.org/10.1002/sce.20264

Shaffer, D. W., & Ruis, A. R. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 175–187) Society for Learning Analytics Research.

Skamp, K. (2007). Conceptual learning in the primary and middle years: The interplay of heads, hearts, and hands-on science. *Teaching Science*, *53*(3), 18–22.

Wei, B. (2020). An exploratory study of teacher development in the implementation of integrated science curriculum. *Research in Science Education*, *50*(6), 2189–2206. https://doi.org/10.1007/s11165-018-9768-x

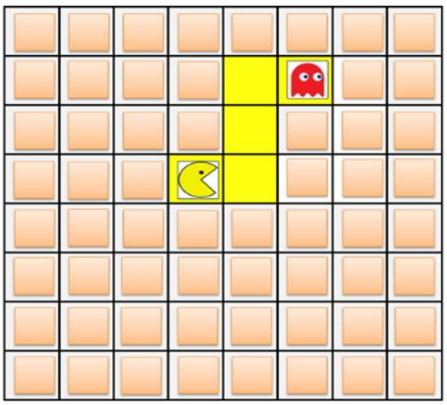
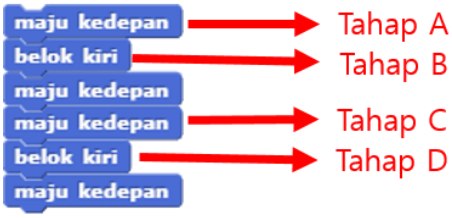
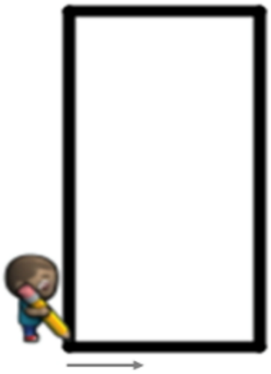
Wiebe, E., Unfried, A., & Faber, M. (2018). The relationship of STEM attitudes and career interest. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(10), 1–17. https://doi.org/10.29333/ejmste/92286

Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015. https://doi.org/10.3102/0034654315626801

APPENDICES

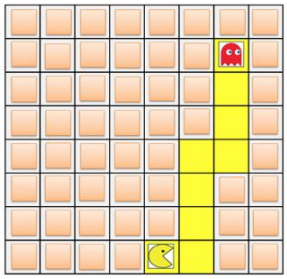
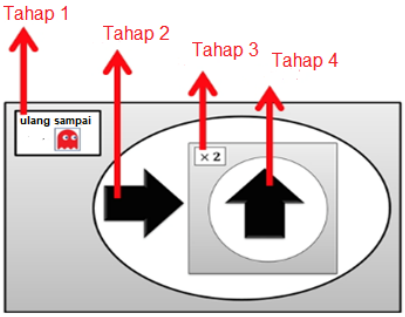
Appendix A: CTA-M (Indonesian Version)

English (Original) Version Cite: Wiebe, E., London, J., Aksit, O., Mott, B. W., Boyer, K. E., & Lester, J. C. (2019, February). Development of a lean computational thinking abilities assessment for middle grades students. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (pp. 456-461). https://doi.org/10.1145/3287324.3287390
English (Original) Version Instrument: https://drive.google.com/file/d/1mKZ8TXSFnVT9NWwnGLP5wHHmTGzqxBVk/view?usp=sharing

Item Number	Item
RG_3 Key = 4	 Perhatikan gambar di atas! Program di bawah ini harus dapat membawa "Pac-Man" ke "hantu" dengan melalui jalur kuning. Tahap manakah dari program di bawah ini yang **salah**?  A. Tahap A B. Tahap B C. Tahap C D. Tahap D
RG_Q7 Key = 1	

<table>
<tr>
<td></td>
<td>Perhatikan gambar di atas! Program di bawah ini harus dapat membuat si pelukis menggambar satu persegi panjang (lebar 50 cm dan panjang 100 cm). Tahap manakah dari program di bawah ini yang salah?

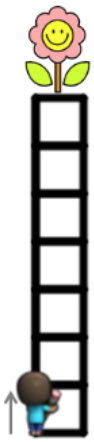

A. Tahap A
B. Tahap B
C. Tahap C
D. Tahap D</td>
</tr>
<tr>
<td>RG_Q11

Key = 3</td>
<td> 
Perhatikan gambar di atas! Program di bawah ini harus dapat membawa “Pac-Man” ke “hantu” dengan melalui jalur kuning. Tahap manakah dari program di bawah ini yang salah?


A. Tahap A
B. Tahap B
C. Tahap C
D. Tahap D</td>
</tr>
</table>

RG_Q12

Key = 1



Program manakah yang si pelukis harus gunakan untuk menggambar tangga yang mencapai Bunga? Ukuran satu anak tangga adalah 30 cm.

A.
```
ulangi sampai Bunga
  ulangi 4 kali
    maju kedepan 30 cm
    belok kanan 90 derajat
  loncat kedepan 30 cm
```

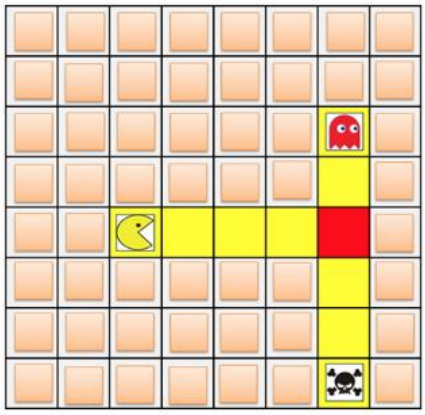
B.
```
ulangi sampai Bunga
  ulangi 4 kali
    maju kedepan 120 cm
    belok kanan 90 derajat
  loncat kedepan 30 cm
```

C.
```
ulangi sampai Bunga
  ulangi 4 kali
    maju kedepan 30 cm
    belok kanan 90 derajat
  loncat kedepan 210 cm
```

D.
```
ulangi sampai Bunga
  ulangi 7 kali
    maju kedepan 30 cm
    belok kanan 90 derajat
  loncat kedepan 30 cm
```

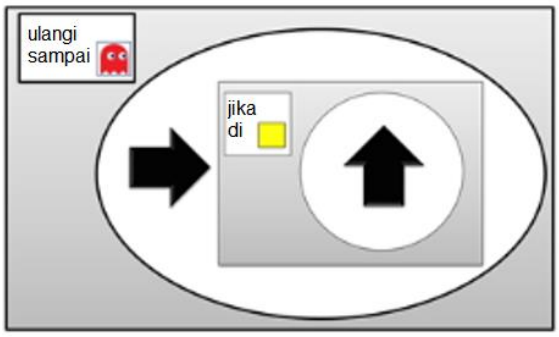
RG_Q13

Key = 2

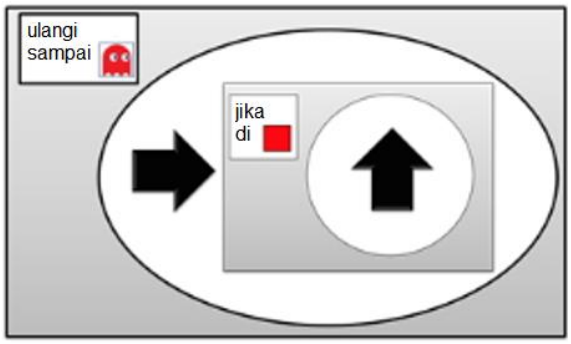


Perhatikan gambar di atas! Program manakah yang dapat membawa “Pac-Man” ke “hantu” dengan melalui jalur kuning?

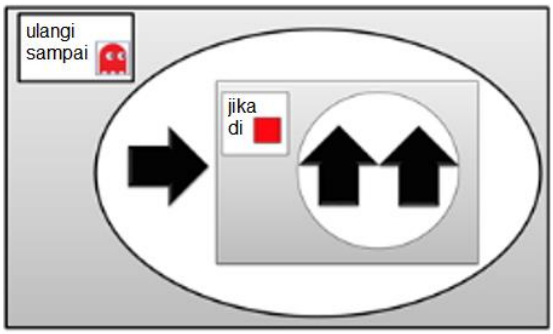
A.

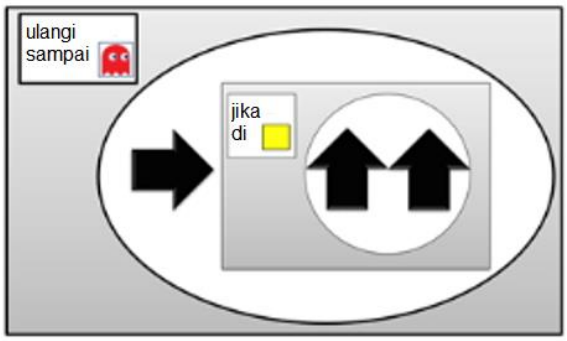
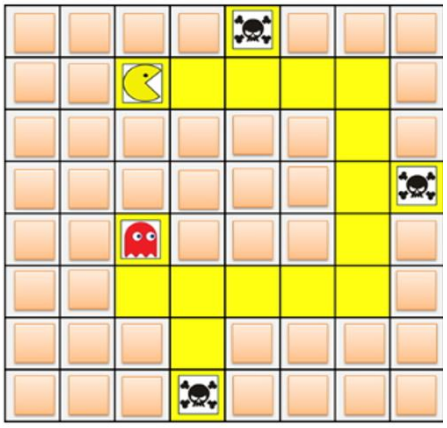
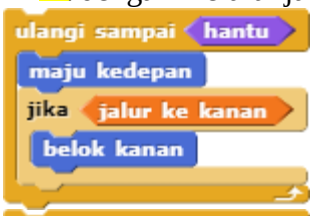


B.

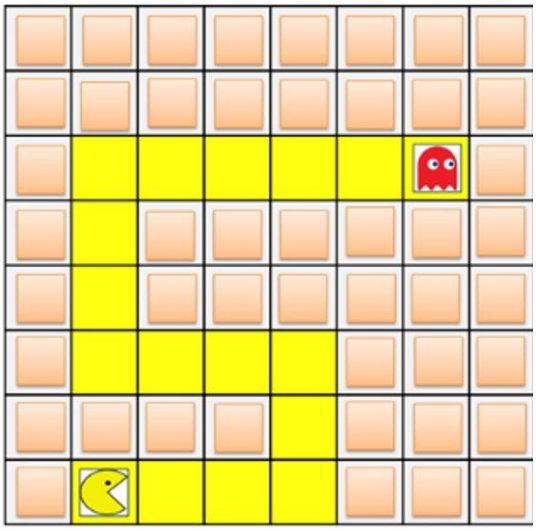


C.



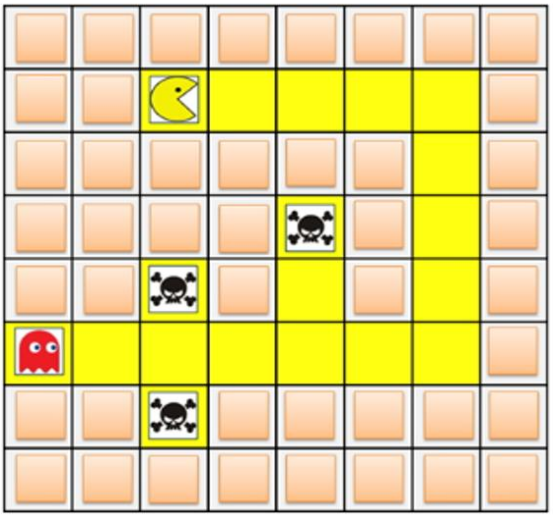
	D. 
RG_Q14 Key = 1	 Perhatikan gambar di atas! Program manakah yang dapat membawa “Pac-Man” ke “hantu” dengan melalui jalur kuning? A.  B.  C.  D. 

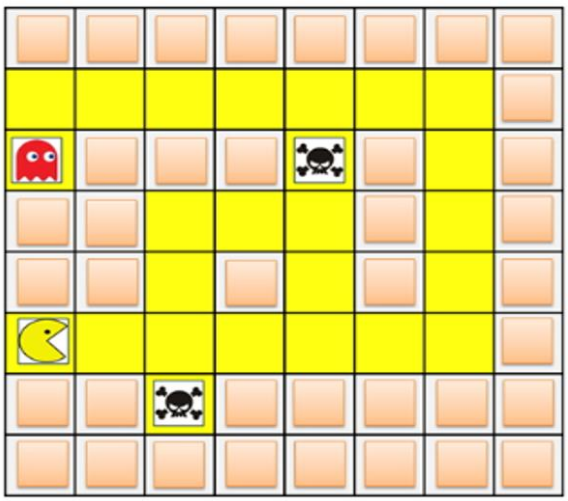
<table>
<tr>
<td>RG_Q16

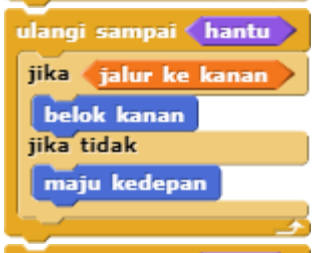
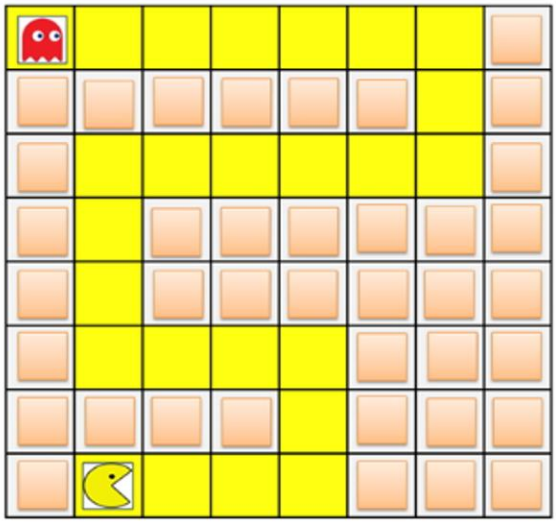
Key = 4</td>
<td>


Perhatikan gambar di atas! Program di bawah ini harus dapat membawa “Pac-Man” ke “hantu” dengan melalui jalur kuning. Tahap manakah dari program di bawah ini yang salah?

A. Tahap A
B. Tahap B
C. Tahap C
D. Tahap D
</td>
</tr>
<tr>
<td>RG_Q17

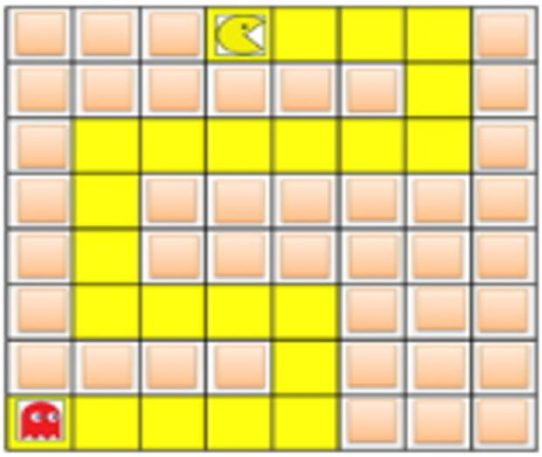
Key = 2</td>
<td>

</td>
</tr>
</table>

	Perhatikan gambar di atas! Program manakah yang dapat membawa “Pac-Man” ke “hantu” dengan melalui jalur kuning? A.  B.  C.  D. 
RG_Q18 Key = 1	 Perhatikan gambar di atas! Program manakah yang dapat membawa “Pac-Man” ke “hantu” dengan melalui jalur kuning?

	A.  B.  C.  D. 
RG_Q19 Key = 2	 Perhatikan gambar di atas! Program di bawah ini harus dapat membawa “Pac-Man” ke “hantu” dengan melalui jalur kuning. Tahap manakah dari program di bawah ini yang **salah**?

<table>
<tr>
<td></td>
<td>


A. Tahap A
B. Tahap B
C. Tahap C
D. Tahap D
</td>
</tr>
<tr>
<td>RG_Q20

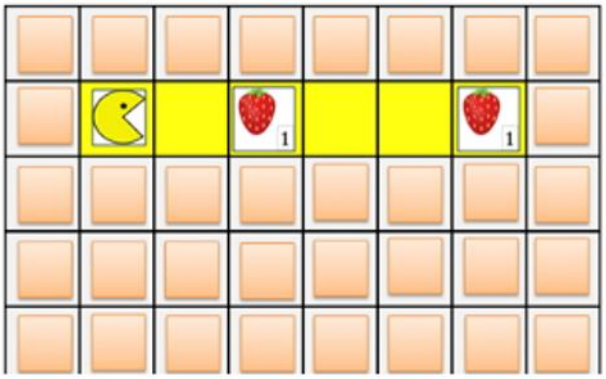
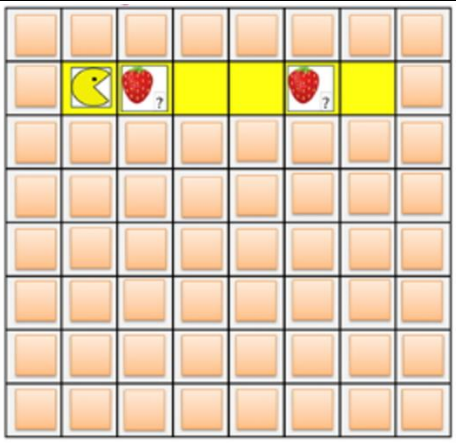
Key = 3</td>
<td>

Perhatikan gambar di atas! Tahap manakah yang hilang () dari program di bawah ini yang dapat membawa “Pac-Man” ke “hantu” dengan melalui jalur kuning?

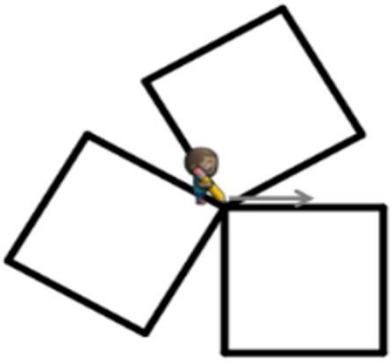

A. 

B. 

C. 

D. Tidak ada tahap yang hilang
</td>
</tr>
</table>

RG_Q23 *Removed for the 23-item instrument Key = 1	 Dengan angka berapakah ????? pada program di bawah ini harus digantikan sehingga "Pac-Man" dapat memakan semua stoberi yang ada di jalur kuning?  A. 1 B. 2 C. 3 D. 5
RG_Q24 Key = 3	 Apakah yang dapat menggantikan bagian ????? pada program di bawah ini sehingga "Pac-Man" dapat memakan semua stoberi yang ada di jalur kuning?  A.  B. 

	C. Stroberi D. bukan stroberi
RG_Q25 Key = 2	Di bawah ini merupakan sebuah rangkaian kode yang membangun suatu program bernama “bungkus.” Program “bungkus” mempunyai fungsi untuk menggambar satu persegi dengan sisi 100 cm. bungkus ulangi 4 kali maju kedepan 100 cm belok kanan 90 derajat Perhatikan gambar di bawah ini! Program manakah yang dapat digunakan oleh si pelukis untuk menggambar desain seperti pada gambar di bawah ini?  A. ulangi 3 kali bungkus belok kanan 120 derajat B. ulangi 3 kali bungkus belok kanan 120 derajat C. ulangi 3 kali bungkus belok kanan 90 derajat D. ulangi 4 kali bungkus belok kanan 90 derajat
RG_Q26 Key = 2	Di bawah ini merupakan sebuah rangkaian kode yang membangun suatu program bernama “bungkus.” Program “bungkus” mempunyai fungsi untuk menggambar satu segitiga dengan sisi 100 cm.

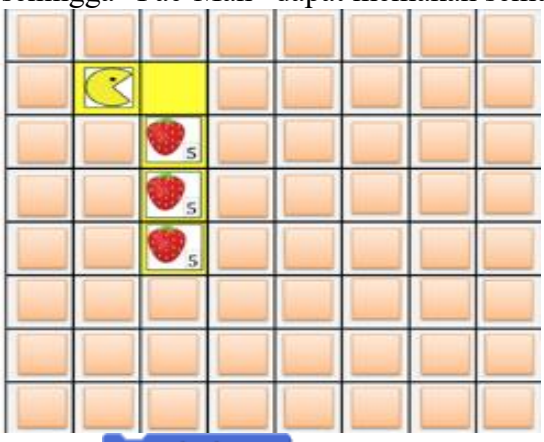
<table>
<tr>
<td></td>
<td>


Perhatikan gambar di bawah ini! Angka berapakah yang dapat menggatikan ????? sehingga si pelukis dapat menggambar desain seperti pada gambar di bawah ini?



A. 15
B. 5
C. 4
D. 3
</td>
</tr>
<tr>
<td>RG_Q27

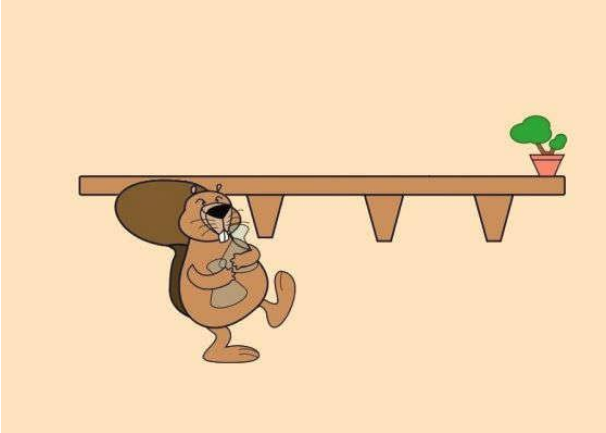
Key = 1</td>
<td>
Di bawah ini merupakan sebuah rangkaian kode yang membangun suatu program bernama “bungkus.” Program “bungkus” mempunyai fungsi untuk mendapatkan dan memakan lima stoberi.

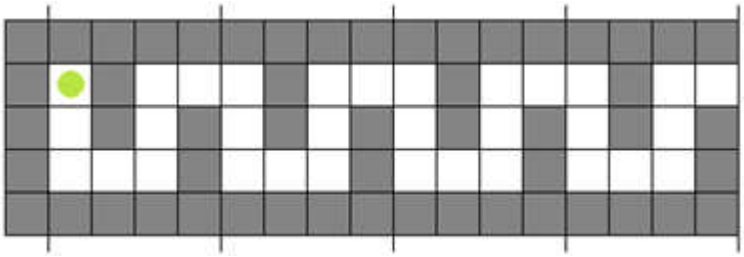
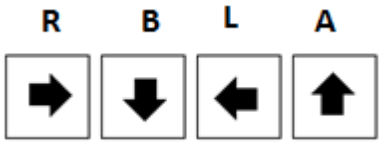
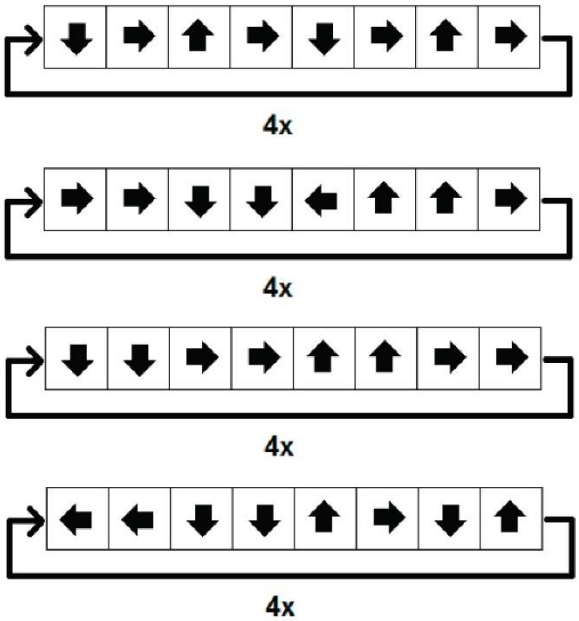

Perhatikan gambar di bawah ini! Program manakah yang dapat digunakan sehingga “Pac-Man” dapat memakan semua stroberi yang ada di jalur kuning?

A.

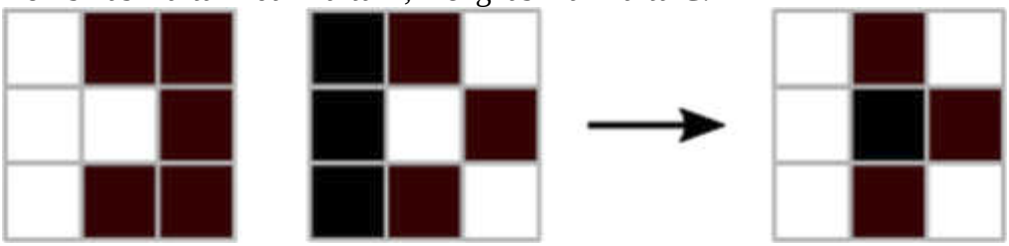
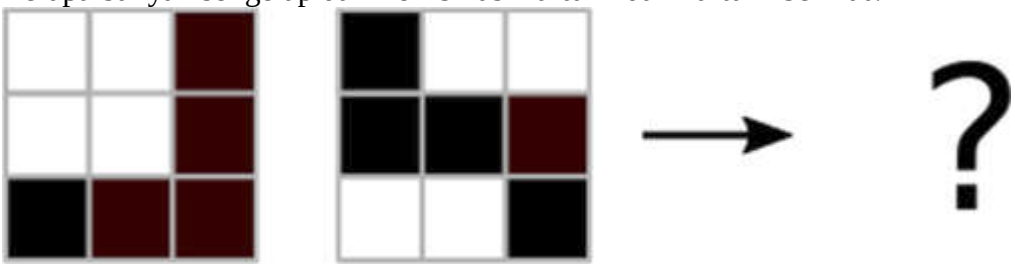
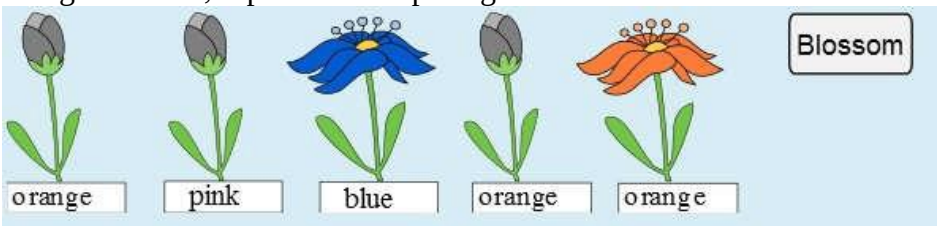
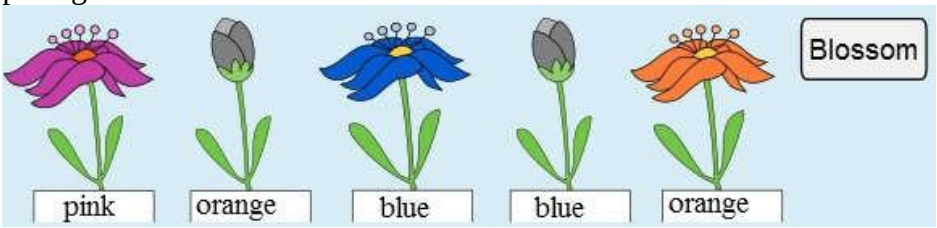

</td>
</tr>
</table>

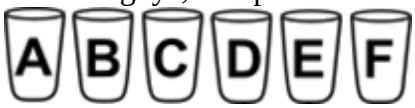
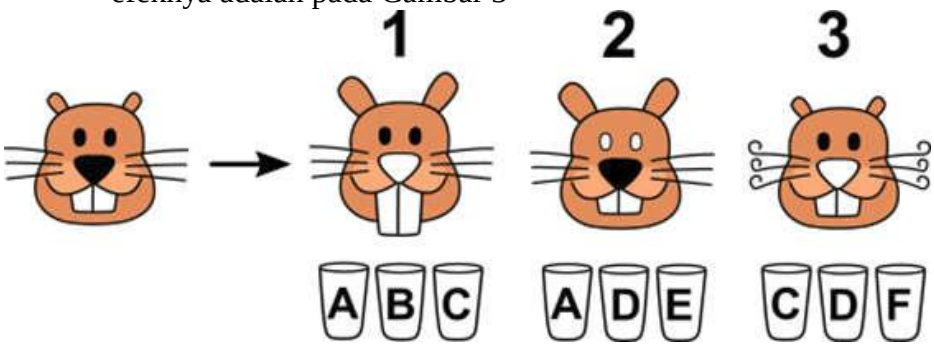
	B.  C. D.

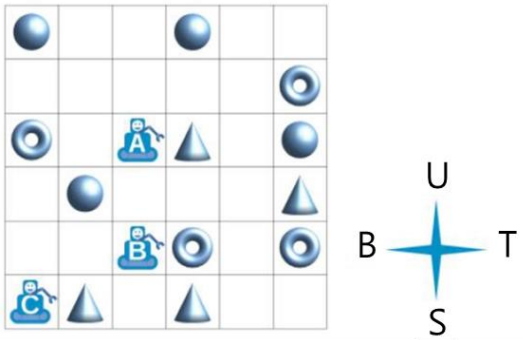
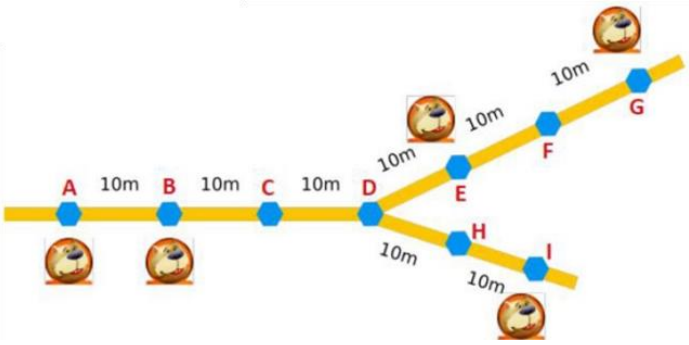
Bebras items

B_Q1 Key = 1	Bebras ingin meletakkan 5 botol di atas rak. Ia ingin agar botol-botol itu dibariskan dari yang paling kecil bagian tengahnya hingga yang terbesar dari kiri ke kanan. Susunlah urutan itu dengan menggeret botol-botol ke atas rak.   Manakah susunan yang paling benar dari yang paling depan ke yang paling belakang? A. Pink, Merah, Kuning, Biru, Hijau B. Merah, Biru, Kuning, Hijau, Pink C. Pink, Kuning, Merah, Biru, Hijau D. Merah, Kuning, Pink, Biru, Hijau
B_Q6	Bantulah si Robot hijau keluar dari lorong.

Key = 3	 Robot dapat bergerak ke empat arah: Kanan (**K**), Bawah (**B**), Kiri (**L**) atau Atas (**A**)  Pilihlah rangkaian arah yang paling benar sehingga si robot hijau dapat keluar dari lorong! 
B_Q7 *Removed for the 23-item instrument Key = 1	Bert memiliki kertas berwarna panjang untuk sebuah pesta. Kertas berwarna ini memiliki tiga warna berbeda (kuning, merah, biru) dengan pola yang berulang secara teratur. Teman Bert, James, telah memotong sebagian kertas, seperti yang ditunjukkan pada gambar di bawah.  James berkata bahwa dia akan mengembalikan kertas yang hilang jika Bert bisa menebak dengan benar ukuran potongan kertas itu. Berapa banyak kotak berwarna yang hilang dari kertas tersebut? A. 31

	B. 32 C. 33 D. 34
B_Q8 Key = 1	Kombinasi kartu A dan kartu B, menghasilkan kartu C.  Berapa banyak sel gelap dari kombinasi kartu D dan kartu E berikut?  A. 3 B. 4 C. 5 D. 6
B_Q9 Key = 2	Jane sedang bermain suatu permainan komputer. Secara rahasia komputer memilih warna-warna untuk lima kuntum bunga. Warna pilihan yang tersedia adalah biru, oranye, dan pink. Pilihan warna-warna itu tidak berubah selama satu permainan. Jane harus menebak warna-warna itu. Lalu Jane menebaknya dan oleh komputer kuntum-kuntum yang warnanya berhasil ditebak segera ditampilkan mengembang, dan yang belum berhasil ditebak tetap ditampilkan sebagai kuntum, seperti terlihat pada gambar berikut.  Itu adalah hasil tebakan pertama. Berikutnya, Jane mendapat kesempatan untuk menebak kedua kalinya, dan hasil tebakannya ditampilkan komputer sebagai pada gambar berikut. 

	Warna-warna apakah yang telah dipilih komputer untuk kuntum-kuntum bunga tersebut? A. pink, biru, biru, oranye, oranye B. pink, biru, biru, pink, oranye C. pink, pink, biru, pink, oranye D. biru, pink, biru, oranye, oranye
B_Q10 Key = 4	Taro si berang-berang menemukan lima jenis ramuan ajaib yang efeknya adalah sebagai berikut: ● Ramuan pertama membuat telinga bertambah panjang ● Ramuan lainnya membuat gigi bertambah panjang ● Ramuan lainnya membuat kumis menjadi keriting ● Ramuan lainnya membuat hidung menjadi putih ● Ramuan terakhir membuat mata menjadi putih Taro menaruh setiap macam ramuan ajaib tersebut dalam sebuah gelas, dan ada sebuah gelas yang berisi air. Keenam gelas tersebut diberi label A sampai dengan F. Malangnya, ia lupa mencatat gelas mana yang mengandung ramuan ajaib apa.  Maka, ia mengadakan percobaan sebagai berikut untuk mengidentifikasi jenis ramuan ajaib pada setiap gelas. ● Percobaan 1: jika ia mengambil ramuan pada gelas A,B dan C, maka efeknya adalah pada Gambar 1 ● Percobaan 2: jika ia mengambil ramuan pada gelas A,D dan E, maka efeknya adalah pada Gambar 2 ● Percobaan 3: jika ia mengambil ramuan pada gelas C, D dan F, maka efeknya adalah pada Gambar 3  Gelas mana yang berisi air? A. A B. B C. C D. D E. E F. F
B_Q13 Key = 2	Di sebuah gudang, ada tiga robot (A, B, C) yang selalu bekerja sama.

	Ketika kelompok robot ini menerima perintah berupa arah mata angin (S, B, T, U), semua robot pada gambar di bawah ini akan bergerak satu kotak pada waktu bersaman dan arah yang diperintahkan. Setelah mengikuti semua daftar perintah, semua robot mendapatkan objek yang ada di kotak terakhir yang mereka lalui. Contohnya, jika kita memberikan daftar perintah U, U, S, S, T kepada kelompok robot itu, robot A akan mendapatkan sebuah corong, robot B akan mendapatkan sebuat cincin, dan robot C akan mendapatkan sebuah corong.  Daftar perintah mana yang dapat membuat kelompok robot itu mendapatkan sebuah bola, sebuah corong, dan sebuah cincin? A. U, T, T, T B. U, T, T, S, T C. U, U, S, T, U D. U, T, T, S, B
B_Q14 Key = 4	Tampak pada peta, lima tempat penginapan berang-berang. Berang-berang ingin menempatkan satu pemberhentian bus di setiap titik biru pada peta. Jarak antara satu titik biru dengan titik biru lainnya adalah 10 meter. Berang-berang memutuskan bahwa jumlah jarak dari penginapan mereka ke pemberhentian bus harus sedekat mungkin.  Dimanakah bus harus berhenti sehingga jarak ke penginapan-penginapanya sekecil mungkin? A. A B. B C. C D. D E. E F. F

	G. G H. H I. I
B_Q15 Key = 2	Agen-agen rahasia Boris dan Berta saling berkomunikasi dengan menggunakan pesan rahasia. Boris ingin mengirim pesan rahasia kepada Berta yang isinya : **MEETBILLYBEAVERAT6** Boris menuliskan setiap karakter pesannya pada grid dengan 4 kolom dimulai dari kiri ke kanan dan dilakukan baris demi baris dimulai dari atas. Boris akan menuliskan karakter X pada tempat di grid yang tidak terpakai. Hasil penulisan pesan Boris dapat dilihat pada gambar di bawah ini.  Selanjutnya, Boris akan membuat pesan rahasia dengan cara menuliskan karakter yang ada di grid mulai dari atas ke bawah dan dilakukan secara kolom demi kolom dimulai dari kolom yang paling kiri, sehingga pesan rahasianya adalah : **MBYVTEIBE6ELERXTLAAX** Berta membalas pesan Boris dengan cara yang sama. Pesan rahasia yang dikirim Berta kepada Boris adalah : **OIERKLTEILH!WBEX** Apakah isi pesan yang ingin dikirimkan Berta kepada Boris ? A. OKWHERETOMEET! B. OKIWILLBETHERE! C. WILLYOUBETHERETOO? A. D. OKIWILLMEETHIM!

Appendix B: Science and CS Attitudes

Science Attitudes Cite: Unfried, A., Faber, M., Stanhope, D. S., & Wiebe, E. (2015). The development and validation of a measure of student attitudes toward science, technology, engineering, and math (S-STEM). *Journal of Psychoeducational Assessment*, *33*(7), 622-639.

No	Item	1D 9 Items				1D 7 Items			
		Estimate	Unweighted MNSQ	Weighted MNSQ	Alpha if item deleted	Estimate	Unweighted MNSQ	Weighted MNSQ	Alpha if item deleted
1	SciAtt_01	-0.070	0.90	0.90	0.810	0.107	1.22	1.22	0.892
2	SciAtt_02	0.218	0.80	0.80	0.799	0.561	1.04	1.03	0.88
3	SciAtt_03	-0.067	0.76	0.76	0.791	0.117	0.89	0.9	0.872
4	SciAtt_04	-0.663	0.66	0.66	0.799	-0.803	0.88	0.88	0.881
5	SciAtt_05	-0.425	0.75	0.76	0.794	-0.436	0.93	0.94	0.877
6	SciAtt_06	-0.147	0.80	0.80	0.800	-0.009	1.02	1.04	0.882
7	SciAtt_07	0.386	1.35	1.33	0.827				
8	SciAtt_08	0.604	2.17	2.16	0.882				
9	SciAtt_09	0.163	0.77	0.76	0.802	0.464	0.93	0.94	0.88
Separation Reliability		0.991				0.99			
EAP/PV RELIABILITY		0.828				0.879			
Cronbach's alpha		0.831				0.896			
Chi-square test of parameter equality		763.04				488.66			
df		8				6			
Sig Level		0				0			
Final Deviance		26512.52378				17685.43533			
Akaike Information Criterion (AIC)		26538.52378				17707.43533			
Akaike Information Criterion Corrected (AICc)		26538.29818				17707.27602			
Bayesian Information Criterion (BIC)		26606.42662				17764.89158			
Total number of estimated parameters		13				11			

S-STEM Science Indonesian Version

SciAtt_01	Saya yakin pada diri saya sendiri ketika saya melakukan aktivitas yang berhubungan dengan sains.
SciAtt_02	Saya akan mempertimbangkan karir di bidang sains.
SciAtt_03	Saya berharap untuk menggunakan sains ketika saya lulus sekolah.
SciAtt_04	Memahami sains akan membantu saya untuk mendapatkan pekerjaan.
SciAtt_05	Saya membutuhkan sains untuk pekerjaan saya di masa depan.
SciAtt_06	Saya tahu saya bisa berhasil di bidang (mata pelajaran) sains.
SciAtt_07	Sains akan menjadi bagian penting dalam pekerjaan saya nanti.
SciAtt_08	Saya dapat medapatkan hasil yang baik di semua mata pelajaran, kecuali sains.
SciAtt_09	Saya yakin saya bisa melakukan pekerjaan lanjutan di bidang sains.

CS Attitudes Cite: Rachmatullah, A., Wiebe, E., Boulden, D., Mott, B., Boyer, K., & Lester, J. (2020). Development and validation of the Computer Science Attitudes Scale for middle school students (MG-CS attitudes). *Computers in Human Behavior Reports*, *2*, 100018. https://doi.org/10.1016/j.chbr.2020.100018

No	Item	1D 9 Items			
		Estimate	Unweighted MNSQ	Weighted MNSQ	Alpha if item deleted
1	CSAtt_01	0.739	1.06	1.05	0.863
2	CSAtt_02	-0.454	0.89	0.88	0.858
3	CSAtt_03	1.734	1.31	1.30	0.874
4	CSAtt_04	-0.331	0.93	0.93	0.856
5	CSAtt_05	-0.647	1.02	1.04	0.858
6	CSAtt_06	-0.660	0.91	0.92	0.856
7	CSAtt_07	-0.939	0.93	0.94	0.862
8	CSAtt_08	0.198	1.00	1.00	0.860
9	CSAtt_09	0.361	1.11	1.11	0.860
Separation Reliability		0.998			
EAP/PV RELIABILITY		0.863			
Chronbach's alpha		0.874			
Chi-square test of parameter equality		2659.46			
df		8			
Sig Level		0			
Final Deviance		25013.36694			
Akaike Information Criterion (AIC)		25039.36694			
Akaike Information Criterion Corrected (AICc)		25039.14167			
Bayesian Information Criterion (BIC)		25107.28874			
Total number of estimated parameters		13			

MG-CSCA Indonesian Version

1	Saya tertarik untuk membuat suatu program komputer baru.
2	Jika saya belajar pemrograman, maka saya dapat meningkatkan kualitas hal-hal yang digunakan orang sehari-hari.
3	Saya mempunyai kemampuan untuk membuat program komputer.
4	Saya tertarik untuk mengetahui bagaimana komputer bekerja.
5	Kemampuan memerogram komputer akan menjadi penting untuk pekerjaan saya di masa depan.
6	Saya ingin tahu bagaimana komputer bekerja.
7	Saya ingin menggunakan kreativitas dan inovasi dalam pekerjaan saya di masa depan.
8	Saat saya menggabungkan matematika dan sains (IPA), saya dapat membuat program komputer yang lebih berguna.
9	Saya yakin saya bisa sukses dalam karir di bidang pemrograman.

Appendix C: Demographics Survey and Exit Ticket

Demographic Information		
Below are a number of questions about your identity. Please take some time to tell us about yourself.		
1	First Name	Open-ended
2	Last Name	Open-ended
2	Which gender do you identify with?	❑ Male ❑ Female ❑ Prefer not to answer
3	What is your current grade level?	❑ 6th grade ❑ 7th grade ❑ 8th grade
4	Which of the following statements BEST describes any previous participation in activities that involve computer science (coding)?	❑ No participation ❑ Low level of participation ❑ Medium level of participation ❑ High level of participation ❑ Expert level of participation
5	Which of the following BEST describes your family economic condition?	● Low economic level ● Medium-low economic level ● Medium-high economic level ● High economic level
6	Do you have any computer at home?	● Yes ● No
7	Do you have any portable computer (e.g., laptop)?	● Yes ● No
8	What is your father's highest level of education?	● Did not graduate from elementary school ● Elementary school ● Middle school ● High school ● College level ● Advanced degree (master's/doctoral)
9	What is your mother's highest level of education?	● Did not graduate from elementary school ● Elementary school ● Middle school ● High school ● College level ● Advanced degree (master's/doctoral)
10	What is your science score on the latest national examinations (elementary)?	Open-ended
11	What is your mathematics score on the latest national examinations (elementary)?	Open-ended

Daily Exit Ticket

No	Question	Option
1	What is your first and last name?	Open Ended
2	Who is your science teacher?	Open Ended
3	What class period is this?	Open Ended
4	Describe what did you learn from today's activity?	Open Ended
5	Please use the following scale to rate how difficult or easy the lesson was today.	● Very easy ● Somewhat easy ● Neither easy nor difficult ● Somewhat difficult ● Very difficult
6	To what extent do you agree with the following statement: "I was able to express my ideas in the model today."	● Strongly disagree ● Disagree ● Neither agree nor disagree ● Agree ● Strongly agree
7	To what extent do you agree with the following statement: "The code I ended the lesson with is my own creation."	● Strongly disagree ● Disagree ● Neither agree nor disagree ● Agree ● Strongly agree

Appendix D: The CISCI Instrument and Experts List

CISCI 1.0 (Indonesian Version)

Item Code	Original Item	Expert_1	Expert_2	Expert_3	Expert_4	Expert_5	Expert_6	Expert_7	Expert_8	Expert_9	Expert_10	Expert_11	Sum of n essential	CVR	Final (Revised after feedback)
CV_01	Beberapa contoh pekerjaan yang menggabungkan sains dan ilmu komputer (koding) adalah astronot, dokter, ilmuwan, pembuat robot, dan insinyur komputer. Saya yakin jenis pekerjaan ini penting.	X	X		X	X	X	X	X	X	X	X	10	0.818	Saya yakin pekerjaan seperti astronot, dokter, ilmuwan, pembuat robot, dan insinyur komputer adalah jenis pekerjaan yang penting karena menggabungkan sains dan ilmu komputer (koding).
CV_02	Saya yakin pekerjaan yang menggabungkan sains dan ilmu komputer (koding/pemerograman) memberikan peluang untuk menghasilkan ide-ide baru.	X	X	X	X	X	X	X	X	X	X	X	11	1	Saya yakin pekerjaan yang menggabungkan sains dan ilmu komputer (koding) memberikan peluang untuk menghasilkan ide-ide baru.
CV_03	Saya yakin pekerjaan yang menggabungkan sains dan ilmu komputer (coding) memberikan peluang untuk berkolaborasi/bekerja sama dengan orang lain.	X	X	X	X	X	X	X	X	X	X		10	0.818	Saya yakin pekerjaan yang menggabungkan sains dan ilmu komputer (koding) memberikan peluang untuk berkolaborasi dengan orang lain.
ASE_01	Saya yakin bahwa saya akan berhasil dalam kegiatan kelas yang menggabungkan sains dan ilmu komputer (koding/pemerograman).	X		X	X	X	X	X	X	X	X	X	10	0.818	Saya yakin bahwa saya dapat memperoleh nilai yang baik dalam pembelajaran yang menggabungkan sains dan ilmu komputer (koding).
ASE_02	Saya yakin dapat mempelajari keterampilan (skills) yang menggabungkan sains dan ilmu komputer (koding/pemerograman).	X	X	X	X	X	X	X	X	X	X	X	11	1	Saya yakin dapat mempelajari keterampilan yang menggabungkan sains dan ilmu komputer (koding).
ASE_03	Saya yakin bisa memahami konsep sains melalui ilmu komputer (koding/pemerograman).	X	X	X	X		X	X	X	X	X		9	0.636	Saya yakin dapat memahami konsep sains dengan memprogram komputer.
CAS_01	Saya yakin bisa bekerja di bidang yang menggabungkan sains dan ilmu komputer (koding/pemerograman).	X	X	X	X	X	X	X	X	X	X	X	11	1	Saya yakin dapat mendapatkan pekerjaan di bidang yang menggabungkan sains dan ilmu komputer (koding).
CAS_02	Saya yakin bisa sukses dalam bidang yang menggabungkan sains dan ilmu komputer (koding/pemerograman).	X	X	X	X	X	X	X	X	X		X	10	0.818	Saya yakin bisa sukses dalam pekerjaan yang menggabungkan sains dan ilmu komputer (koding).
CAS_03	Saya yakin saya dapat menggabungkan sains dan ilmu komputer (koding/pemerograman) dalam karir masa depan saya.	X	X	X		X	X	X	X	X	X	X	10	0.818	Saya yakin saya dapat menggabungkan sains dan ilmu komputer (koding) dalam karir masa depan saya.
CIN_01	Saya menikmati belajar sains dengan menggunakan koding atau pemerograman.	X		X	X		X	X	X	X		X	8	0.455	Saya yakin bahwa menggabungkan pengetahuan sains dan ilmu komputer akan berguna setelah saya lulus sekolah.
CIN_02	Saya berniat untuk mendapakan pekerjaan yang menggabungkan sains dan ilmu komputer (koding/pemerograman).	X	X	X	X	X	X	X	X	X	X	X	11	1	Saya berkeiningan untuk mendapakan pekerjaan yang menggabungkan sains dan ilmu komputer (koding).
CIN_03	Saya sangat suka berbicara dengan orang-orang yang menggabungkan sains dan ilmu komputer (koding/pemerograman) dalam pekerjaan mereka.	X	X	X	X	X	X	X		X	X	X	10	0.818	Saya senang berdiskusi dengan orang-orang yang menggabungkan sains dan ilmu komputer (koding) dalam pekerjaan mereka.
PSR_01	Saya membahas pekerjaan yang menggabungkan sains dan ilmu komputer (koding/pemerograman) dengan orang tua saya.	X	X	X	X	X	X	X		X	X	X	10	0.818	Saya akan mendiskusikan pekerjaan yang menggabungkan sains dan ilmu komputer (koding) dengan orang tua saya sebagai cita cita saya.
PSR_02	Orang tua saya yakin saya bisa sukses dalam pekerjaan yang menggabungkan sains dan ilmu komputer (koding/pemerograman).	X	X		X	X	X	X	X	X	X	X	10	0.818	Saya yakin orang tua saya akan mendukung jika di masa depan saya memilih bekerja di bidang yang menggabungkan sains dan ilmu komputer
PSR_03	Saya mengenal seseorang di keluarga saya yang menggabungkan sains dan ilmu komputer (koding/pemerograman) dalam pekerjaannya..	X	X	X		X	X	X		X		X	8	0.455	Saya mengenal seseorang di keluarga atau diluar keluarga saya yang menggabungkan sains dan ilmu komputer (koding) dalam pekerjaannya.
PSR_04	Saya mengenal seseorang yang menggabungkan sains dan ilmu komputer (koding/pemerograman) dalam pekerjaannya.		X	X	X	X	X	X	X	X	X	X	10	0.818	Saya terinspirasi/termotivasi untuk berkarir di bidang yang menggabungkan sains dan ilmu komputer karena pengaruh orang yang saya kenal yang bekerja di bidang tersebut
PSR_05	Saya memiliki panutan (role model) yang menggunakan ilmu komputre/koding/pemerograman dalam karir sainsnya.	X	X	X	X	X	X	X	X	X	X	X	11	1	Saya menjadikan orang yang saya kenal yang bekerja di bidang sains dan ilmu komputer sebagai panutan

Expert List

No/Code	Expert Name	Highest Level of Education	Occupation
1	Expert 1	M.Ed.	Science Education Researcher/Ph.D. Student
2	Expert 2	M.Ed.	High School Biology Teacher
3	Expert 3	M.Ed.	Science Education Researcher

4	Expert 4	Ph.D.	Science Education Researcher
5	Expert 5	M.Ed.	High School Biology Teacher
6	Expert 6	B.Ed.	Middle School Science Teacher
7	Expert 7	Ph.D.	Biology Education Faculty
8	Expert 8	Ph.D.	Biology Education Faculty
9	Expert 9	Ph.D.	Science Education Faculty
10	Expert 10	Ph.D.	Chemistry Education Faculty
11	Expert 11	M.Ed.	Physics Education Researcher/Ph.D. Student

CISCI 2.0 (Indonesian and English Versions)

Construct	Code	Revised/Final Indonesian	Revised/Final English
CIS Career Value	CV_01	Saya yakin pekerjaan seperti astronot, dokter, ilmuwan, pembuat robot, dan insinyur komputer adalah jenis pekerjaan yang penting karena menggabungkan sains dan ilmu komputer (koding).	I believe jobs like astronaut, doctor, scientist, roboticist, and computer engineer are important because they combine science and computer science (coding)
	CV_02	Saya yakin pekerjaan yang menggabungkan sains dan ilmu komputer (koding) memberikan peluang untuk menghasilkan ide-ide baru.	I believe jobs that combine science and computer science (coding) help me come up with new ideas.
	CV_03	Saya yakin pekerjaan yang menggabungkan sains dan ilmu komputer (koding) memberikan peluang untuk berkolaborasi dengan orang lain.	I believe jobs that combine science and computer science (coding) provide opportunities to collaborate with others.
CIS Academic Self-efficacy	ASE_01	Saya yakin bahwa saya dapat memperoleh nilai yang baik dalam pembelajaran yang menggabungkan sains dan ilmu komputer (koding).	I am confident that I will do well in classroom activities that combine science and computer science (coding).
	ASE_02	Saya yakin dapat mempelajari keterampilan yang menggabungkan sains dan ilmu komputer (koding).	I am confident that I can learn skills that combine science and computer science (coding).
	ASE_03	Saya yakin dapat memahami konsep sains dengan memprogram komputer.	I am confident that I can understand science through computer science methods (coding)
CIS Career Self-efficacy	CAS_01	Saya yakin dapat mendapatkan pekerjaan di bidang yang menggabungkan sains dan ilmu komputer (koding).	I am confident I can work in a field that combines science and computer science (coding).
	CAS_02	Saya yakin bisa sukses dalam pekerjaan yang menggabungkan sains dan ilmu komputer (koding).	I am confident that I can be successful in a field that combines science and computer science (coding)
	CAS_03	Saya yakin saya dapat menggabungkan sains dan ilmu komputer (koding) dalam karir masa depan saya.	I am confident I can combine science and computer science (coding) in my future career.
CIS Career Interest and Goal	CIN_01	Saya yakin bahwa menggabungkan pengetahuan sains dan ilmu komputer akan berguna setelah saya lulus sekolah.	I believe that combining scientific concepts and computer science would be useful after I am done with schooling.
	CIN_02	Saya berkeiningan untuk mendapakan pekerjaan yang menggabungkan sains dan ilmu komputer (koding).	I want to enter a career that combines science and computer science (coding).
	CIN_03	Saya senang berdiskusi dengan orang-orang yang menggabungkan sains dan ilmu komputer (koding) dalam pekerjaan mereka.	I would feel comfortable talking to people who combine science and computer science (coding) in their work.
CIS Parental Supports and Role Models	PSR_01	Saya akan mendiskusikan pekerjaan yang menggabungkan sains dan ilmu komputer (koding) dengan orang tua saya sebagai cita cita saya.	I would talk to my parents about my goal to have a job that combines science and computer science (coding).
	PSR_02	Saya yakin orang tua saya akan mendukung jika di masa depan saya memilih bekerja di bidang yang menggabungkan sains dan ilmu komputer	I believe that my parents would support me in the future if I choose a career that combines science and computer science (coding)
	PSR_05	Saya menjadikan orang yang saya kenal yang bekerja di bidang sains dan ilmu komputer sebagai panutan	I have role models who use coding in their science careers.

Appendix E: T-STEM Science and CT Instruments (Self-Efficacy Only)

T-STEM Science Scale

Item Code	Item	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
PSTEB_1	I am continually improving my science teaching practice.					
PSTEB_2	I know the steps necessary to teach science effectively.					
PSTEB_3	I am confident that I can explain to students why science experiments work.					
PSTEB_4	I am confident that I can teach science effectively.					
PSTEB_6	I understand science concepts well enough to be effective in teaching science.					
PSTEB_7	Given a choice, I would invite a colleague to evaluate my science teaching.					
PSTEB_8	I am confident that I can answer students' science questions.					
PSTEB_9	When a student has difficulty understanding a science concept, I am confident that I know how to help the student understand it better.					
PSTEB_10	When teaching science, I am confident enough to welcome student questions.					
PSTEB_11	I know what to do to increase student interest in science.					

T-STEM CT Cite: Boulden, D. C., Rachmatullah, A., Oliver, K. M., & Wiebe, E. (2021). Measuring in-service teacher self-efficacy for teaching computational thinking: development and validation of the T-STEM CT. *Education and Information Technologies*, 1-27.

DIRECTIONS:

For each of the following statements, please indicate the degree to which you agree or disagree.
Even though some statements are very similar, please answer each statement. There are no "right" or "wrong" answers. The only correct responses are those that are true for you. Whenever possible, let the things that have happened to you help make your choice.

"The participants were asked to apply the following definition of CT as they responded to each item:
A problem-solving process that applies key ideas from computer science such as algorithms, abstraction, pattern recognition, and decomposition. Computational thinking may involve programming and computers but does not have to. It is a human thought process that utilizes computational tools and concepts to solve problems.

Computational Thinking Teaching Efficacy and Beliefs
Directions: Please respond to these questions regarding your feelings about ***your own*** teaching.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1. I am continually improving my computational thinking teaching practice.	○	○	○	○	○
2. I know the steps necessary to teach computational thinking effectively.	○	○	○	○	○
3. I am confident that I can explain to students why computational thinking experiments work.	○	○	○	○	○
4. I am confident that I can teach computational thinking effectively.	○	○	○	○	○
5. I wonder if I have the necessary skills to teach computational thinking.	○	○	○	○	○
6. I understand engineering concepts well enough to be effective in teaching computational thinking.	○	○	○	○	○
7. Given a choice, I would invite a colleague to evaluate my computational thinking teaching.	○	○	○	○	○
8. I am confident that I can answer students' computational thinking questions.	○	○	○	○	○
9. When a student has difficulty understanding a computational thinking concept, I am confident that I know how to help the student understand it better.	○	○	○	○	○
10. When teaching computational thinking, I am confident enough to welcome student questions.	○	○	○	○	○
11. I know what to do to increase student interest in computational thinking.	○	○	○	○	○

Appendix F: Participants per School

No	School (Pseudonym)	Teacher (Pseudonym)	Grade	Total	All			Consented and Assented			Completion (3 attendances and worksheets)		Survey/ Incomplete
					Computational	Drawing	Neither	Computational	Drawing	Neither	Computational	Drawing	
1	School A	Regina	7	191	128	63	0	92	50	0	42	13	87
			9	68	57	11	0	50	11	0	20	4	37
		Tamara	7	64	18	46	0	14	36	0	9	15	26
2	School B	Wilda	7	89	56	33	0	41	27	0	26	14	28
3	School C	Alishba	7	74	50	24	0	42	16	0	33	5	20
			8	50	0	0	50	0	0	17	0	0	17
			9	55	0	0	55	0	0	11	0	0	11
4	School D	Shirly	7	146	146	0	0	90	0	0	50	0	40
		Dina	8	126	0	0	126	0	0	13	0	0	13
			9	96	0	0	96	0	0	37	0	0	37
		Aiden	8	4	4	0	0	2	0	0	-	-	-
			9	35	35	0	0	23	0	0	-	-	-
5	School E	Samuel	7	152	0	152	0	0	101	0	0	35	66
6	School F	Vivian	7	131	43	4	84	18	2	18	16	1	21
7	School G	Avery	7	128	128	0	0	77	0	0	48	0	29
8	School H	Brighita	7	86	0	86	0	0	71	0	0	27	44
		Natalie	7	88	0	88	0	0	55	0	0	18	37
		David	7	28	0	28	0	0	19	0	0	13	6
Total				**1611**	**665**	**535**	**411**	**449**	**388**	**96**	**244**	**145**	**519**
										Without Grade 9	224	141	
										Total Used in Intervention study	**365**		

Appendix G: Protocol for Back-Translation Process

Protocol

Even though there are at least four different procedures (i.e., back-translation, bilingual technique, committee approach, and pretest; Brislin, 1970) to translate a research instrument from one language to another language, Jones et al. (2001) and Cha et al. (2007) recommended using at least a combination of two methods. We will combine the back-translation and pretest methods to translate the instruments used in our study. According to Brislin (1970), the back-translation method starts with a bilingual expert translating the instrument from the original language to the target language, in the context of this study is from English to Bahasa. Another bilingual expert will then blindly translate the translated instrument back to the original language (Bahasa to English). Afterward, an English native will compare the instrument translated from Bahasa to English with the original English version to assess the accuracy of the words and sentences as well as the intended meaning. Any differences and inaccuracies will be discussed with the bilingual experts until a consensus is reached.

Next, the instrument (Bahasa version) will be piloted to the targeted participants. This piloting is intended to identify any potential problems that may occur when the instrument is used in the larger study (Brislin, 1970; Cha et al., 2007). However, before piloting the instrument, two teachers familiar with the content and participating students will go through the instrument to identify any wording problems or complex sentences. Revisions will be made based on the teachers' suggestions. The revised version will then be used in the pilot study. The pilot's data will be analyzed using Item Response Theory (IRT)-Rasch to gather the instrument's psychometric properties. We will use these psychometric properties to remove problematic items. The final version used in the larger study will be the instrument after problematic items have been removed.

References

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185-216. https://doi.org/10.1177/135910457000100301

Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: issues and techniques. *Journal of Advanced Nursing*, *58*(4), 386-395. https://doi.org/10.1111/j.1365-2648.2007.04242.x

Jones, P. S., Lee, J. W., Phillips, L. R., Zhang, X. E., & Jaceldo, K. B. (2001). An adaptation of Brislin's translation model for cross-cultural research. *Nursing Research*, *50*(5), 300-304. https://journals.lww.com/nursingresearchonline/Abstract/2001/09000/An_Adaptation_of_Brislin_s_Translation_Model_for.8.aspx