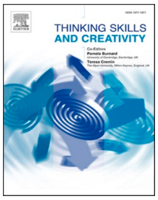
Thinking Skills and Creativity 38 (2020) 100742

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: www.elsevier.com/locate/tsc



Investigating the effects of a systematic and model-based design of computer-supported argument visualization on critical thinking

Berty Nsolly Ngajie [a], Yan Li [a,*], Dawit Tibebu Tiruneh [b], Mengmeng Cheng [c]

[a] College of Education, Zhejiang University, Hangzhou, PR China
[b] Faculty of Education, East China Normal University, Shanghai, PR China
[c] School of Education, Hangzhou Normal University, Hangzhou, PR China

ARTICLE INFO

Keywords:
Critical thinking instruction
Computer-supported argument visualization
Infusion approach
Higher education
First principles of instruction

ABSTRACT

Computer-supported argument visualization (CSAV) is an educational technology that has been identified in many studies to be effective in stimulating university students' critical thinking (CT) skills. However, the instructional design principles and strategies that are effective in enhancing students' CT with CSAV have not been sufficiently explored. This study investigated the effectiveness of a systematic and model-based design of CSAV-based CT instruction on the enhancement of university students' CT skills. Two groups of students (experimental and comparison) were recruited at a Chinese university to participate in a preparatory course, Modern Education Technology (MET). The experimental group, which comprised 39 students, was fully engaged with CSAV cooperative activities designed systematically based on the First Principles of Instruction model. The comparison group, which comprised 31 students received a regular instruction designed by the subject teacher. Participants in both groups were pre- and post-tested with the Halpern Critical Thinking Assessment (HCTA) test. Findings disclosed that participants in the experimental group significantly outperformed the comparison group in overall CT performance, and especially in hypothesis testing, problem-solving, and argument analysis sub-skills of the HCTA test. This result suggests that a systematic and model-based approach of designing CSAV-based CT instruction is effective in developing university students' CT skills. Implications of the findings for designing CT-supportive subject-matter instruction in line with the CSAV strategy are discussed.

1. Introduction

Improving university students' critical thinking (CT) has been a major goal of higher education for several decades. It is more so in the 21 st century (Facione, 2020; Halpern, 2014). Employers of the era of artificial intelligence want students to develop CT skills needed in the modern workplace (Bezanilla, Fernández-nogueira, Poblete, & Galindo-domínguez, 2019; Koc, Kahn, Koncz, Salvadge, & Longenberger, 2019). Individuals who are critical thinkers are known to gather relevant information and reach well-reasoned conclusions, make accurate predictions and decisions, assess the credibility of sources, identify cause-effect relationships, and communicate effectively with others (Halpern, 2014). The ability to think critically is generally linked with improved decision-making when faced with real-life problems, and with a tendency to become a more active and informed citizen (Dwyer, Hogan, & Stewart, 2012).

---

* Corresponding author at: College of Education, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, 310058, PR China.
E-mail address: yanli@zju.edu.cn (Y. Li).

https://doi.org/10.1016/j.tsc.2020.100742
Received 3 August 2020; Received in revised form 12 September 2020; Accepted 26 September 2020

The amount of information available to university students in this information age is massive. It is assumed that CT instruction through active instructional methods allows students to supersede the simple memorization of information by acquiring an intricate comprehension of the information while studying and in the future workplace (Halpern, 2014). Previous research on CT instruction has focused more on teaching domain-general CT skills required in everyday life independently from domain-specific courses. However,the issue of integrating CT instruction within domain-specific courses and enhancing domain-specific CT skills has been under-researched (Fischer et al., 2014; Tiruneh, Verburgh, & Elen, 2014). There are a few studies that focused on integrating CT skills as part of the standard curriculum (Swartz & Mcguinness, 2014; Tiruneh, Mieke, Spector, Gu, & Elen, 2017). Nevertheless, the issue of identifying which instructional approaches are more effective in fostering CT skills in subject-matter instruction remains inadequately addressed.

Computer-supported argument visualization (CSAV) is a technology-based pedagogical tool which provides a learning environment suitable for enhancing complex learning outcomes, such as CT (Davies, 2011). CSAV has been empirically tested and found useful in stimulating college students' CT more than other CT instructional strategies (Davies, 2009; Harrell & Wetzel, 2015; Hitchcock, 2017; Twardy, 2004; Van Gelder, 2015). Research has shown that courses that focus on CSAV-based CT instruction produce more significant CT improvements compared to those that use traditional stand-alone undergraduate CT courses (Hitchcock, 2017). However, instructional design principles and strategies that are most effective in enhancing students' CT with CSAV have not been sufficiently researched. One approach to research on instructional effectiveness is to express what learning processes determine the effectiveness of an instructional method (Clark & Mayer, 2011). It is not sufficient to attribute CT gains to CSAV; it is also essential to know how it works by specifying the instructional design processes that determine the effectiveness of the CSAV strategy.

It is contended in the present study that efforts to stimulate subject-matter CT using CSAV infusion may benefit from a systematic and model-based approach of designing subject-matter instruction. Considering the potent role of university students' CT abilities in academics and future workplace, and the positive results indicated in the literature that the CSAV strategy improves university students' CT skills, it is crucial to explore the instructional design principles and strategies that work best with the CSAV strategy. This will help eliminate the doubts surrounding this promising strategy. This study examined the effects of a systematic and model-based design of CSAV on the development of CT skills. The remainder of this paper presents the theoretical and empirical background of the study. It then presents the purpose of the study and research hypotheses followed by the method and findings sections.

2. Theoretical and empirical background

2.1. Approaches to teaching CT

There are two views on teaching CT emerging from the longstanding domain-general versus domain-specific debate (Davies, 2006, 2013; Ennis, 1990; McPeck, 1990; Moore, 2004). The former advocates that CT skills are best taught as stand-alone courses so that domain-specific content instruction does not overshadow them. The latter, on the other hand, emphasizes that CT skills are best developed in specific domains because thinking skills that are necessary for one field may not be the same as those needed in other fields (McPeck, 1990). From this debate, Ennis (1989) identified four approaches to teaching CT, which include General, Infusion, Immersion, and Mixed approaches. The Immersion approach is described as implicit method to CT instruction, while the General and Infusion approaches are explicit methods. Many studies have shown that explicit approaches to CT instruction result in more CT gains than the implicit ones (Abrami et al., 2008; Alan Bensley & Spero, 2014; Bensley, 2010; Marin & Halpern, 2011).

The General approach explicitly teaches CT skills independently without the use of subject matter content. The Infusion approach, which is the focus of this study, incorporates CT instruction in a regular subject while making general CT principles explicit (Ennis, 1989; Swartz & Mcguinness, 2014). No matter the theoretical tradition from which CT originated, most of the early CT teaching programs mainly adopted the General approach (Swartz & Mcguinness, 2014). However, this approach became less effective in improving students' CT skills both in their other course work and beyond school life (Bailin, 2002). Also, given the need for in-depth competency in domain-specific content knowledge, the transition in recent years entails the teaching of CT skills as part of the standard curriculum using embedded approaches such as the Infusion approach (Bailin, 2002; Swartz & Mcguinness, 2014; Tiruneh, De Cock, & Elen, 2018).

2.2. The Infusion approach and its instructional strategies

The Infusion approach has its roots in the sociocultural theory, which states that active and effective learning takes place in meaningful interaction as learners learn to think individually by first reasoning with others (Mercer, Hennessy, & Warwick, 2019). Adopting the Infusion approach helps teachers plan the existing curriculum contents and a range of thinking skills together without sacrificing the learning of subject matter contents (Bensley, 2010; Dewey & Bento, 2009; McGuinness, 2000; Swartz & Mcguinness, 2014). As noted by Bensley (2010), specific CT skills are targeted while making CT criteria, rules, and methods explicit through guided training in the form of exercises. The exercises are focused on skills assessment, and feedback is provided on practice. To achieve this, it uses several explicit strategies to help students develop CT skills. Some frequently used strategies have been reviewed, revealing that dialogue, authentic or anchored instruction, and mentoring (Coaching/tutoring/modeling) have been found to have positive effects on students' CT skills (Abrami et al., 2015; Tiruneh et al., 2014).

Abrami et al. (2015) concluded in their meta-analysis that dialogue in small cooperative learning groups, student exposure to problems that are relevant to them, which include examples and mentoring are those strategies that have positive effects on students' CT skills. Nonetheless, the meta-analysis did not focus specifically on undergraduate CT instruction (David Hitchcock, 2017).

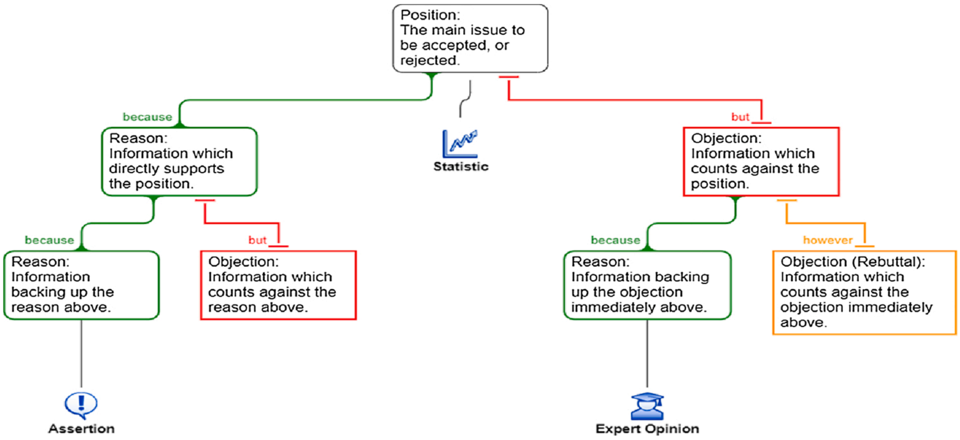


Fig. 1. Scaffolding reasoning in Rationale including assertions, statistic, and expert opinion as evidence (editor page in *Rationale*).

Moreover, CSAV, which is reported by many studies to improve university students' CT (Davies, 2009, 2012; Dwyer, Hogan, & Stewart, 2013; Eftekhari, Sotoudehnama, & Marandi, 2016; Van Gelder, 2015) is neither included in the meta-analysis nor in Tiruneh et al. 's (Tiruneh et al., 2014) systematic review.

2.3. CSAV-based CT instruction in the infusion approach

CSAV is another strategy the Infusion approach uses to make CT skills explicit in subject-matter instruction (McGuinness, 2000; Sun, Wang, & Wegerif, 2019; Swartz & Mcguinness, 2014). It is an ICT-based pedagogical tool that provides an environment where teachers help students to practice reasoning by providing guidance, scaffolding, and feedback (Van Gelder, 2007). As a cognitive mapping tool that uses dedicated computer programs to represent inferences between premises and conclusions, it enables the teaching and understanding of CT skills. These specialized computer programs help learners to express complex ideas while engaging in high-order thinking tasks via visual presentation of relevant information (Dwyer et al., 2013; Sun et al., 2019). Sweller and colleagues' work cited in Dwyer et al. (2013) on cognitive load suggests that presenting information in "visual-verbal" format (diagrams and text) decreases cognitive load and enhances learning. CSAV's ability to support both diagrams and text enhances the hierarchical organization of information, which in turn optimizes the way information is organized both in the "working memory" and "long-term memory" thus, helping to promote CT skills (Dwyer et al., 2012).

CSAV tools, like the *Rationale* software adopted in this study, make use of boxes to represent propositions, while arrows represent inferential relationships between the various propositions (Dwyer et al., 2013). Boxes are coded with different colors indicating the type of proposition. At the same time, arrows are labeled using words or phrases (called "inference indicators") to scaffold the type of relationship that exists between the propositions. The color distinction helps learners to identify the structure and logic of instances of reasoning quickly. These standard reasoning pointers include words like "because," and "as a result"; objection indicator words like "but", "even though", "however", etc. Fig. 1 depicts scaffolding reasoning in *Rationale*. Representing information in this way has been found to reduce cognitive load, thereby facilitating memory and comprehension (Dwyer et al., 2012).

As mentioned earlier, research shows that the CSAV strategy improves CT skills. Working with students majoring in EFL in the Iranian context, Eftekhari et al. (2016) compared CT skills development using the *Rationale* CSAV software with the traditional paper and pencil format. The results disclosed that students who used the CSAV performed significantly higher than those who used the conventional pencil and paper in general CT as well as in the sub-CT skills of inference and inductive reasoning assessed using the California Critical Thinking Skills Test (CCTST).

Similarly, research conducted by Hitchcock (2017) found that courses that focused on CSAV-based CT instruction produced more significant improvements than those that used the traditional stand-alone undergraduate CT courses and computer-assisted instruction combined with writing instructions. Also, Davies' (2009) research with an Australian Economic History course students using the *Rationale* CSAV tool found that students felt quite satisfied with the CSAV technique at post-test. They attested that it was beneficial in understanding learning tasks requiring assessment, the nature of arguments, and CT. They also found it helpful in summarizing scientific papers, analyzing academic arguments as well as in determining academic arguments. Therefore, they suggested that it should also be adopted in the other courses.

Except for the research conducted by Tiruneh et al. (2017), a majority of the CT-infused studies with and without CSAV in higher education mainly seem to teach domain-general CT skills independently from domain-specific courses. Although researchers do use examples related to the courses in which the interventions were conducted, they do not demonstrate how the whole course was designed and delivered from chapter to chapter or topic to topic (Alan Bensley & Spero, 2014; Angeli & Valanides, 2009; Dwyer, Hogan, & Stewart, 2011; Eftekhari et al., 2016). For instance, in Davies (2009), it is indicated that the CSAV methodology on how to extract the structure of arguments from simple to complex prose-texts was given to students. However, how the instruction was designed was not indicated. Alan Bensley and Spero (2014) simply made general CT principles explicit without actually designing for

CT infusion. Similarly, the study by Eftekhari et al. (2016) does not establish how the writing course contents were combined with the CT skills to enhance students' CT skills. Although students in the CSAV and pencil and paper groups were systematically instructed on how they could extract the structure of an argument, the instructional materials used in the intervention were not directly derived from the course textbook.

The assessment of CT in subject matter content is often done using standardized tests that are used to assess general CT skills. Some of the results show that the subject matter CT instruction can increase achievement in domain-general CT skills, while others argue the contrary. For instance, Tiruneh et al. (2017) conducted an intervention in which they explicitly emphasized domain-specific CT (Physics) skills. A domain-general critical thinking assessment test, the Halpern Critical Thinking Assessment (HCTA), was used to verify the achievement of domain-general critical thinking skills. The results did not show significant improvement in domain-general critical thinking skills. It is not clear how subject matter contents are designed to be assessed with a domain-general CT test like HCTA. It is also not known whether greater achievement in a domain-general test signifies a deep understanding of the domain-specific subject matter in which these tests are used. We argue that there needs to be a correspondence between the CT skills taught in subject-matter contents and those measured by any CT assessment instrument. Accordingly, the methodology of the HCTA test was adapted to design subject-matter CT activities in this study.

In the Chinese context, CT education is also one of the areas targeted in ongoing educational reforms in universities. The Secretary-General emphasized at the national education conference held in Beijing, on September 10, 2018, that grades and diplomas should not be the sole form of evaluating students in higher education (Chen & Yu, 2018). Yanhong and Lin (2014) added that thinking skills should also be included in the evaluation profile. However, there are relatively few studies on CT, and most of the available research is mainly translations of foreign works or exploring the concept of CT itself (Huan & Hu, 2014). Although some CT courses such as the ongoing MOOCs at Zhejiang University[1] and Nanjing Forestry University[2] have recently been set up for the explicit teaching of CT, research indicates that the teaching of CT in subject matter content is described as unclear (Huan & Hu, 2014; Chen & Yu, 2018; Zhang, 2017). Some reasons why CT is not given due consideration among several others include large class sizes, heavy class schedules, and a lack of teacher training in CT instruction (Tian & Low, 2011, 2012). There is, therefore, the need for in-depth research with respect to research methods and implementation of CT instruction in the Chinese context. It is based on this backdrop that this study was conducted in an interdisciplinary teacher education course, MET with preservice primary school teachers of Chinese Language and Culture. We consider teacher educators as being in the frontline with teachers who have the responsibility of teaching students how to think critically. As future primary school teachers, they should be provided with a good number of opportunities to allow them to have the critical thinker's experience themselves while still at university so that they may be able and willing to transmit it to their students.

3. Purpose of the study and research hypotheses

The purpose of this study was to examine the effect of a systematic and model-based design of CSAV-Infused subject matter CT instruction on the development of CT skills. We argue in this paper that CT includes several complex skills and dispositions which require not just a single strategy, but a combination of explicit teaching strategies. A single strategy may not guarantee the ultimate goal of improving learners' CT skills. This study hypothesizes that CSAV combined with other explicit strategies in a well-coordinated skill-approach may lead to significant improvement in CT acquisition. Strategies such as "a course instructor trained in anchored instruction," "student cooperative learning groups engaging with problems within subject-matter instruction," and "coaching" are incorporated.

Considering the need for explicit instructional strategies for the teaching of undergraduate CT in subject-matter instruction, the CSAV-infusion in this study adopted Ennis' (1989) Infusion approach to CT instruction. The Infusion approach with CSAV in this study assumed that CSAV must be combined with several explicit instructional strategies to teach selected CT skills alongside regular subject-matter content. It allows the illustration of general principles of thinking and dispositional skills that may be applicable to domain-specific CT problems.

Bearing in mind the need to understand how the CSAV strategy improves CT, this contribution attempts to demonstrate how CSAV can be applied in the Infusion Approach. Based on previous studies

(Hitchcock, 2017; Rapanta & Walton, 2016), CSAV-based CT instruction is neglected as an effective CT instructional strategy. One possible explanation could be because research that examines the instructional design principles for subject-matter CSAV-based CT infusion has not been conducted.

Merrill's (2002) First Principles of Instruction model features among the several instructional design models known to provide clear guidance when designing instruction geared towards learners' constructivist acquisition of complex knowledge and skills. It promotes the usage of valid and contextually appropriate learning exercises for the teaching of the subject matter. After a review of various instructional design models and theories, Merrill (2002) identified prescriptive principles that are commonly used in the various instructional designs that promote learning: problem centered, activation, demonstration, application, and integration principles.

It has been argued that these empirically tested and validated instructional design principles emerging from the teaching of subject-matter offer specific guiding principles on how to design learning environments that promote CT development (Merrill, 2002; Tiruneh et al., 2017). At the time of this study, it is not known whether this model has been used to design a CSAV-based CT instruction. The model was adopted in this study because it takes a systematic approach and has been empirically tested in promoting complex learning

---

[1] https://www.icourse163.org/course/ZJU-1206354803

[2] https://www.icourse163.org/course/NJFU-1001755007

Table 1
The targeted CT skills, subject matter competence and guiding questions for instruction.

The five CT sub-skills and their definitions	Subject-matter competence	Sample guiding questions for instruction
Verbal reasoning: These skills aid one to recognize and avoid unconvincing evidence regardless of the context.	Identify ambiguity of terms; Recognize errors in poorly designed teaching and learning materials; Interpret the outcome of MET-based lesson plans.	What are the strengths of …?; what are the implications of …?; How does…affect…?; How could …be used to …?
Thinking as hypothesis testing: These skills allow one to observe and formulate experiments and assess the validity of a given hypothesis.	Identify the choice of multimedia resources that promote learning; Inspect the adequacy of observations to draw conclusions *e.g., from your observation, what can you infer(conclude)about downloading images, videos, ppts, pdfs when preparing for a multimedia-based lesson?* Check if samples are adequate in size as well as any bias possible when generalizing.	What would happen if …; If …what would be the…; How does… affect…?; How does this … ties with what we learned before? Do we have a good reason to make a generalization, a conclusion, an inference?
Argument analysis: These skills allow one to expound on and examine the reliability and strengths of one's arguments and those of others.	Recognize the main parts of the argument on issues related to modern education technologies; Deduce a precise statement for a given data set; Criticize the validity of a generalization inferred from some teaching experiments. Provide different perspectives of a given problem	What is the problem or solution …and reasons supporting it? Could you provide a counter-argument (objection) for… ? ; What is the counter-argument for…?
Likelihood and uncertainty Analysis: These skills are useful in predicting probabilities of successes and failures in day to day decision-making and problem-solving regardless of context.	Predict the probability of events e.g., what can you predict ; Recognize assumptions that have to be maintained in a generalization drawn from the results of teaching experiments; Recognize that more information is needed when making a decision; Make effective predictions.	How can you predict that the choice of multimedia resources will result in successful or unsuccessful learning outcomes? ; What other information is needed to make a decision about…?;
Problem-solving and decision making: These skills allow the identification and creation of new and alternate ways of making decisions and solving problems.	Identify several alternative ways of solving problems; Examine importance procedures needed when solving teaching and learning with technology problems; Evaluate solutions related to Educational Technology problems; Make decisions based on evidence; Use the right analogies in solving problems related to Education Technology	What is…analogous to?; Do you have sufficient evidence to make a sound decision? What alternative method is suitable for solving this kind of problem?

like CT (Lo & Hew, 2017; Tiruneh et al., 2017). This study also employed the systematic and model-based approach proposed by Tiruneh et al. (2017). It is described as "an approach to design learning environments based on comprehensive instructional design model oriented towards the promotion of complex learning" (Tiruneh et al., 2017, p. 4). Consequently, CSAV tasks in this contribution were designed following Merrill's First Principles of Instruction using the subject-matter contents of the MET course.

The following hypotheses (H) were formulated in the study:

H1. The combination of explicit instructional strategies in a systematic and model-based course design increases students overall CT performance.

H2. The systematic design of cooperative learning activities when implementing a CSAV tool based on the principles of the cognitive test items results in significant improvement on sub-CT skills.

H3. Students' engagement with the CSAV tool improves their argument analysis skills.

4. Methodology

4.1. Participants

The participants of this study were third-year Chinese Language and Culture majors from a relatively big public university in the city of Hangzhou, China. Two groups of junior students were recruited to participate in this quasi-experimental study. Their ages ranged between 19 and 22 years old. Both groups were all native speakers of Mandarin Chinese. The experimental class comprised of 48 students, while the comparison group consisted of 45 students. In the experimental group, while 48 students (44 females and 4 males) took part in the HCTA test at pre-test, only 39 students (37 females and 2 males) were able to complete the post-test. And in the comparison group, 45 students (41 females and 4 males) took part in the HCTA pre-test, but only 31 (28 females and 3 males) completed the post-test. Given that the post-test was administered during the week that students were preparing for the end of semester exams, the data of 22 students were missing. Some either had just one of the test formats at post-test or did not complete all the items in one of the test. So they were completely excluded in the entire experiment.

4.2. Design, development and implementation of instructional activities

The intervention focused on an undergraduate course, titled Modern Education Technology (MET). The first step in the design process was to choose a CT definition, and the skill sets that students were to be assessed on using the HCTA test. The teaching of CT is

Table 2
The general structure of lessons based on Merrill’s "First Principles of Instruction"

Stages in the lesson	Principles	Instructor and learners’ activities in CSAVIN
Introduction	Activation of prior knowledge and problem presentation	The instructor introduced the objectives of each lesson and those of the targeted thinking skills selected for the lesson. For instance, during the first intervention lesson, after introducing the lesson’s objectives, She then listed Halpern’s thinking skills described in Table 1. Next, she explained the meaning of each of the five skills, their importance, and how they can be beneficial in learning about Education Technologies. Finally, the instructor presented the problem, and students worked in their cooperative learning groups.
Presentation of contents and providing learners with thinking vocabulary	Demonstration of active and skill-full thinking	Based on students’ responses, before the instructor calls up one of the group members to share their ideas with the whole class, the instructor explicitly models, examples of CT skills on how to use language, knowledge, and skills in completing a given task. Members in other groups and or the instructor may provide feedback in the form of support in case help or correction is needed or summarizes the ideas of various groups when they present a variety of ideas.
Helping students think about their thinking	Application of new knowledge	The instructor poses thoughtful questions that guide students to reflect on their thinking. For example, she may ask questions related to any of the five skills. Using likelihood and uncertainty, for example, *What would happen if such poorly designed PPT is used; could you explain how you predicted that the videos in the problem would be helpful to David?* This enables the instructor to know if they are doing it right or wrong and how to give corrections. The instructor also provides feedback and encourages them to practice more with the *Rationale* CSAV software by providing incomplete visual maps and also using simple to complex tasks to make their thing visible. This helps to engage them in purposeful reflection.
Conclusion	Integration for the future transfer	Here, the instructor provided daily life tasks to be executed in order to integrate the lesson into students’ lives and to encourage future transfer. E.g… a news journal article for students to analyze the information contained in it about micro-lectures in China using the *Rationale* CSAV tool. Also, students were required to summarize the CT skills that they used to analyze the articles each time they completed a task.

based on the assumption that there exist identifiable and teachable thinking skills, which, when students learn to recognize and apply, could become better thinkers (Halpern, 2014, p. 19). The conception of CT by Halpern as "…the use of those cognitive skills or strategies that increase the probability of a desirable outcome" (Halpern, 2014, p.8) was selected as it captures major skills that could be taught explicitly. The skills include "verbal reasoning," "argument analysis," "hypothesis testing," "likelihood and uncertainty," and "problem-solving, and decision-making" skills. This definition was chosen because it captures the main concepts that are included in most CT descriptions. It allows for the teaching of CT using a skills approach.

Next, new topics from the curriculum that were to be used to practice the five CT skills were determined. The intervention mainly focused on five topics, (1) obtaining and using ICT-based multimedia instructional resources; 2) PowerPoint (PPT) production; 3) production of micro-lectures and quick webpage design systems (tools); 4) smart classroom environment; 5) gamification and horizon report). A description of the targeted CT skills, subject matter competence, and guiding questions for instruction are presented in Table 1.

The third step was to design and develop the instructional activities based on the First Principles of Instruction model (Merrill, 2002). The course instructor (fourth author), and the rest of the authors collaborated in designing the experimental lessons. It was ensured that the targeted CT skills were included as part of subject-matter activities in the design and development process of the activities. Several activities were designed that students had to complete in class, not as take-home assignments.

The intervention was implemented during the 2018/2019 school year over seven sessions, and every session lasted 2 h.

4.3. Training of instructor

Given that the same course instructor was to teach the experimental and comparison groups, she was well-coached by the first author to ensure that she implemented the experimental lessons as per the design. The instructor was well informed about the goal of the teaching intervention and was briefed on what was expected of her at every given point. Her opinion was sought concerning the draft version of each intervention lesson, and she provided valuable feedback. Discussions continued regularly either through emails and face-to-face video calls or text messages whenever she had doubts, or when the first author wished to ensure that the instructor has fully understood some of the lesson activities.

4.4. The CSAVIN condition

Participants in the experimental group, referred to as the CSAV Infusion (CSAVIN) group, were divided into cooperative learning groups of 4 or 5 students per group. Prior to lesson implementation, *Rationale* accounts were created for students to ease their access and use of the platform. Participants were informed about the objectives of the study at the beginning of the first session of the intervention. In addition to instructor explanations, they were also provided with videos guiding them on how to create thinking maps

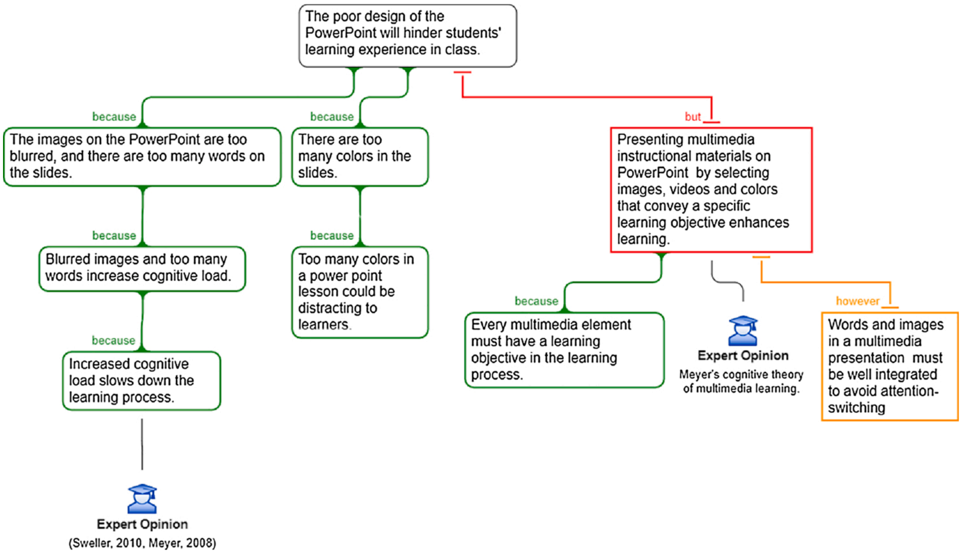


Fig. 2. An example of a reasoning evaluating a lesson designed in Power Point selected from course material.

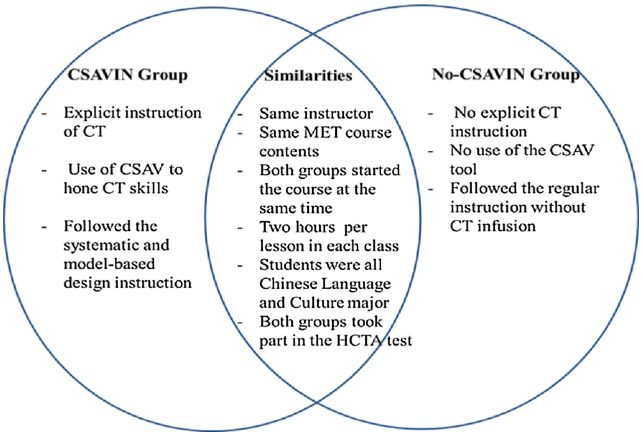


Fig. 3. Similarities and differences between the CSAVIN and the No-CSAVIN Groups.

with *Rationale*. Additional instructions were made available on Power Points (PPT) during every activity. Table 2 describes the general structure of lessons designed and implemented in the order of Merrill's "First Principles of Instruction."

In applying the First Principles of Instruction model during the implementation of the instructional activities, it was ensured that: 1) students engaged in solving MET course relevant tasks; 2) the teacher activated students' previous knowledge based on the selected MET topics; 3) the instructor demonstrated new knowledge and skills by providing examples, coaching and feedback to students; and 4) students were given more opportunities to apply and integrate the newly learned knowledge and skills using both incomplete and entirely new tasks to engage in their cooperative learning groups as they hone their thinking skills during interaction. At this stage, they were also provided with lots of exercises to practice analyzing MET problems with the use of the CSAV tool, *Rationale.* Fig. 2 is an example of a reasoning map evaluating a Power Point lesson selected from course material. The opportunity allowed them to make their thinking visible and encouraged them to develop the disposition to think critically. It was ensured that learners were mentored with clear instructions towards the achievement of subject-matter outcomes in CT in different activities per topic. The first author supervised the implementation of all seven sessions to ensure that lessons were implemented as designed. Students were attended to wherever help was needed either by peers or by the teacher.

4.4.1. The No-CSAVIN condition

The comparison group referred to as the No CSAV Infusion (No-CSAVIN) group, followed the regular instruction designed and implemented by the course instructor. There was no mention of anything about CT or CT skills. To have a general overview of the instructional procedure in the comparison group, the first author also attended all the sessions of the No-CSAVIN group. The instructor began each lesson by presenting the general objectives to be achieved at the end of the lesson. Then she presented the topic with some

Table 3
Sample HCTA item and subject-matter learning activity on an argument analysis task.

Sample HCTA Item	Sample subject-matter learning activity
Some universities are considering a new requirement that every student must do some meaningful public service to graduate. *In no more than five sentences, explain your position on this proposal.*	After exploring how the sections of "group map rotation" and "technical information" are realized in the link below https://genzhe-cmm-xuewangye.kuaizhan.com *In no more than five sentences, explain the piece of advice you would give a novice web designer.*

oral questions. The instructor did not make use of thinking vocabulary to provide explanations for their answers unlike in the experimental class. The problems and instructions were not explicit; and they did not make use of the CSAV tool.

Except for the explicit teaching of the CT skills and the use of the CSAV tool *Rationale* for the experimental group, the two groups were comparable in other issues as shown in Fig. 3 below.

4.5. Data collection and analysis

4.5.1. The CSAV tool, Rationale

The CSAV tool adopted in this study was *Rationale.*· It was used as an additional strategy for dispositional purposes in the last part of lessons to practice the five CT skills measured by the HCTA test. *Rationale* is a great tool that helps users to learn about the fundamentals of good reasoning and CT[3] . It allows them to create and manipulate a visual representation showing the logical flow of the argument structure (Davies, Barnett, & Gelder, 2019). Still, it is most suitable for topics involving CT, which is an essential part of all fields, including the field of Educational Technology. Empirical evidence demonstrates that training in CT skills with *Rationale* yield significant improvement (Davies, 2009; Eftekhari et al., 2016; Van Gelder, 2015). Therefore, it was hypothesized that student engagement with the CSAV tool will result to significant improvement in their argument analysis skill.

4.5.2. HCTA test

The HCTA test is a general CT skill test that was used in this study to assess subject-matter CT skills. The majority of the highly used tests, including Watson-Glaser Critical Thinking Appraisal (WGCTA) and CCTST, are mainly multiple-choice tests (Ku, 2009). As opposed to these tests, the HCTA test uses the multiple-choice as well as the constructed response test formats to assess the skills and dispositional components of CT (Ku, 2009). It assesses a set of five CT skills that were targeted in this study. Altogether there are 20 items based on a variety of real-life problems such as health, education, politics, and social policy. Each scenario is made up of the two kinds of assessment formats; the open-ended response (constructed-response items) and a short-list of answers to select from (forced-choice items). Previous studies indicated that the test's internal consistency ranges between .79–.88 (Dwyer et al., 2012). The internal consistency of the test in this study was accepted based on Cronbach's $\alpha = .78$ and .81 for the pre-test and post-test of both formats combined.

The principles of formulating the test the items was adopted in designing domain-specific learning activities based on MET contents for students to discuss in cooperative learning groups and to execute with the CSAV tool, as presented in Table 3. It was hypothesized that using the principles of formulating the HCTA items in subject-matter instruction would facilitate transfer to general CT skills scenarios assessed by the HCTA test. Each task included at least two or more CT skills to be developed during each session.

The two groups completed the paper version of the HCTA test in session one at the pre-test, and the last session was reserved purposely for the post-test. The procedure of administering the HCTA was carefully followed; thus, participants first completed the open-ended response items, and the forced-choice items were completed last. The test was administered in an ecological classroom learning environment with both formats lasting 50 and 75 min, respectively. Some students still could not finish on time. They had to send it by email after the class. Students were awarded candies, chocolate, and some writing material in return for their participation. Only the experimental group (CSAVIN) followed the designed activities. The comparison group (No-CSAVIN) took part only in the pre-test and post-test of the HCTA test.

At the end of the intervention, an interview was conducted with six students from the experimental group to capture their perceptions of the design of the intervention on their CT skills and dispositions. The interview constituted of seven open-ended questions (See Appendix A). Q1−3 verified the changes students observed in the teaching compare to the previous sessions before the intervention and how group discussion changed the way they think about thinking. Q4−5 verified the effects of group discussion and the use of CSAV on their thinking. Q6−7 verified whether or not students perceived that the intervention helped them learn the five CT skills of hypothesis testing, argument analysis, verbal reasoning, and problem-solving, likelihood.

5. Results

An analysis of covariance (ANCOVA) was used to test H1, while paired sample *t*-tests were used to test H2 and H3. The effect size that was decided upfront for each of the hypotheses was $p < .05$.

---

[3] ReasoningLab.com, https://www.rationaleonline

5.1. H1: the combination of explicit instructional strategies in a systematic and model-based design increases students' overall CT performance

To test H1, an ANCOVA was conducted to assess whether there was a significant difference between the comparison group and the experimental group in the HCTA Test on overall CT performance.

Table 4 presents the means and standard deviations for the comparison group and the experimental group on the HCTA test before and after controlling for the pre-test scores. As is evident from the table, virtually no difference between the No-CSAVIN group and the CSAVIN group remains after differences on the HCTA pre-test were controlled. The results of the ANCOVA indicated that after controlling for the pre-test scores, there is a significant difference between the No-CSAVIN and CSAVIN groups in the HCTA test scores on the post-test $F\ (167) = 9.66$, $P <.003$. The analysis of the covariance for CT achievement using the systematic and model-based design as a function and the HCTA's post-test scores as a covariate, both groups scored higher at post-testing compared with pre-test CGP (pre-test: $M = 96.61$, $SD = 12.69$; post-test: $M = 100.61$, $SD = 13.09$) and EGP (pre-test: $M = 94.05$, $SD = 15.61$; post-test: $M = 105.64$, $SD = 12.62$). However, participants in the experimental group scored higher overall on CT, as accounted for by post-testing differences.

5.2. H2: the systematic design of cooperative learning activities when implementing a CSAV tool based on the principles of the cognitive test items result in significant effects on sub-CT skills

To test H2, a paired sample *t*-test was conducted to compare the means between groups and within groups. Given that the *t*-test is very robust, it revealed no statistical difference between the No-CSAVIN and CSAVIN groups at the pre-test of the five CT sub-skills. An inspection of CT sub-skills revealed that there was a significant effect on all the sub-skills in the CSAVIN group compared to the No-CSAVIN group. Within groups, the CSAVIN group scored significantly higher on the post-test compared with pre-testing of the five CT sub-skills of hypothesis testing $t\ (1,\ 38) = -4.07$, $p <.\ 001$, verbal reasoning ($t\ (38) = -3.51$, $p <.001$, argument analysis $t\ (38) = -2.72$, $p <.010$, likelihood/uncertainty ($t\ (38) = -3.03$, $p = .004$, and problem-solving $t\ (38) = -4.41$, $p <\ 001$. Nonetheless, compared to pre-test performance, the No-CSAVIN group performed better in verbal reasoning $t\ (31) = -2.84$, $p < .008$ and argument analysis $t\ (31) = -2.38$, $p < .024$ on the post-test.

As displayed in Table 5 , it was observed that the CSAVIN group's post-test means and standard deviations of the five CT sub-skills were significantly higher than those of the No-CSAVIN group especially in hypothesis testing (CSAVIN: post-test: $M = 22.28$, $SD = 4.09$; No-CSAVIN: $M = 19.00$, $SD = 4.70$); problem-solving (CSAVIN: $M = 33.46$, $SD = 3.99$; No-CSAVIN: $M = 31.16$; $SD = 4.98$); and argument analysis (CSAVIN: $M = 24.18$, $SD = 6.05$; No-CSAVIN: $M = 24.16$; $SD = 4.21$). However, both groups were almost equal on likelihood and uncertainty (CSAVIN: $M = 15.97$ $SD = 2.62$; No-CSAVIN: $M = 16.00$, $SD = 2.22$). Compared to pre-test results, the comparison group performed better in verbal reasoning (CSAVIN: $M = 9.74$, $SD = 2.07$; No-CSAVIN: $M = 10.29$, $SD = 2.49$). However, the improvement in the CSAVIN group was greater at post test compared to pretest (1.48) unlike in the No-CSAVIN group (0.87).

5.3. H3: Students' engagement with the CSAV tool improves their argument analysis skills

A paired sample *t*-test was again conducted on the pre-test and post-test scores of the argument analysis skill to examine the effect produced by engagement with the CSAV tool in the CSAVIN group. The means and standard deviation obtained from the HCTA test revealed a significant improvement ($t\ (38) = -2.72$, $P < .010$) at post-test ($M = 24.18$; SD = 6.05) compared to pre-test ($M = 21.46$; $SD = 6.11$). Table 6 presents the descriptive statistics of CSAVIN students' performance on the argument analysis skill.

5.4. Students' perception towards the intervention

An interview was conducted with six students (S1, S2, S3, S4, S5 and S6) two days after the intervention. The interview took place two days after the intervention. Instead of having to interview students one by one, which is often affected by interviewer interaction (Lin, 2018), the investigator created a *WeChat* group where each student posted all their responses of the seven interview questions. The data obtained from the interviews showed that the students were affirmative toward the intervention lessons on their thinking in terms of engaging in cooperative learning groups and using CSAV, which created in them the disposition to think, especially with regards to completing a task which in turn enhanced their CT skills Students perceivedindicated that the infusionfact that the instructor made the CT skills explicit during lessons had effects on their thinking over timeencouraged them to engage in critical thinking behaviours such as identifying assumptions, making accurate predictions, and assessing the credibility of sources. The full findings are discussed in the discussion section to fulfill the purpose of providing explanations for the quantitative data results in the previous section.

Table 4

Adjusted and Unadjusted group means and standard deviations using the HCTA's pre-test scores as a covariate.

		Unadjusted		Adjusted	
Groups	*N*	*M*	*SD*	*M*	*SD*
No-CSAVIN Group	31	94.05	12.69	100.61	13.09
CSAVIN Group	39	96.61	15.61	105.64	12.62
Total	70	95.15	14.34	103.41	12.98

Table 5
Means and standard deviations for CT performance in the five skills by groups.

		Pre-test	Post-test		
The five CT skills	*N*	*M (SD)*	*M (SD)*	*t*	*p*
Hypothesis Testing					
CSAVIN	39	20.00 (4.73)	22.28 (4.09)	−4.08	.001*
No-CSAVIN	31	18.87 (3.98)	19.00 (4.70)	−.163	.872
Verbal Reasoning					
CSAVIN	39	8.26 (2.88)	9.74 (2.07)	−3.51	.001*
No-CSAVIN	31	9.42 (2.08)	10.29 (2.49)	−2.84	.008*
Argument Analysis					
CSAVIN	39	21.46 (6.11)	24.18 (6.05)	−2.72	.010*
No-CSAVIN	31	22.52 (4.63)	24.16 (4.21)	−2.38	.024*
Likelihood/Uncertainty					
CSAVIN	39	14.79 (2.98)	15.97 (2.62)	−3.04	.004*
No-CSAVIN	31	15.52 (2.32)	16.00 (2.22)	−1.11	.277
Problem-solving					
CSAVIN	39	29.54 (5.54)	33.46 (3.99)	−4.41	.001*
No-CSAVIN	31	30.29 (5.04)	31.16 (4.98)	−1.26	.219

Note. *$P < .05$

6. Discussion

This study tested three hypotheses. H1: The combination of explicit instructional strategies in a systematic and model-based design increases students' CT performance. H2: The systematic design of cooperative learning activities when implementing a CSAV tool based on the principles of the cognitive test items results in significant effects on sub-CT skills. H3: Students' engagement with the CSAV tool improves their argument analysis skills.

In H1, findings showed that combining explicit instructional strategies in a systematic and model-based design increases students' overall CT performance. Training a course instructor in anchored instruction, having students practice CT skills in cooperative learning groups, and using the CSAV tool to build dispositional skills towards CT in a systematic and model-based design is a promising approach to designing CSAV-based CT instruction.

It was observed that both groups performed better at the post-test in overall CT performance. The reasons for the improvement in CT skills in the No-CSAVIN group may be attributed firstly to maturity over the semester, and test effects, in this case, the HCTA test. This interpretation is in line with (Marsden & Torgerson, 2012), who argue that irrespective of any intervention, students mature over time as they go through university. Secondly, the fact that the same test was repeated could trigger students to remember questions or create some learning awareness leading to better performance during the second testing time. Thirdly, the MET course itself already involves a lot of problem-solving. Using regular instruction without the explicit teaching of CT in the No-CSAVIN group, the course instructor might implicitly teach students some CT strategies without telling them, and students might have learned how to solve MET problems without even knowing they were using CT skills.

However, the much greater improvement observed in the experimental group could be ascribed to the systematic and model-based design of the infusion activities, which were based on the principles of the cognitive test items. This result is in line with the results obtained in students' CT assessments with the HCTA test and student responses in the after-intervention interview. The results suggest that guiding collaborative learning groups around authentic problems, training a subject-matter instructor, practicing visualizing thinking with a CSAV tool, as well as requiring students to summarize CT skills after completing tasks allows them to master the skills they learned. This finding is in line with the suggestions put forward in the meta-analysis conducted by Abrami et al. (2015), who indicated that a combination of explicit CT instructional strategies could result in more significant gains in CT than when a single strategy is used alone. This analysis is supported by students' interview responses. They perceived that the instructional intervention had effects on their thinking over time. In Q1−2 of the interview all the six students interviewed observed that there was a sudden change from learning alone to a lot of group discussion and collaboration on hands-on-task which enriched their engagement in thinking from different perspectives unlike in the first part of the course before the intervention. They also pointed out that feedback on problems was provided to them immediately. 5 of the interviewees except 1 reported that this strategy improved their learning than when they were working alone. S3 for instance explained that:

> *To solve a problem in the process of thinking, I will consider the issue from different aspects to answer, from the pros and cons to think about the problem. In the process of thinking together in a group, everyone is free to speak and exchange ideas with other members. I think*

Table 6
Descriptive statistics of performance in the argument analysis skill.

Argument analysis	Pre-test	Post-test		
	M(SD)	*M (SD)*	*t*	*p*
	21.46 (6.11)	24.18(6.05)	−2.72	.010

Note. *$P < .05$; CSAVIN group: 39.

> *the thinking experience is a gradual process. I also think group learning is better than individual learning. In the process of group learning, each person expresses his or her own views, so that he or she will have multiple perspectives on a problem and solve the problem more thoroughly and profoundly. This effect cannot be achieved by learning individually.*

This confirmed the findings of past studies that found that opportunities to dialogues (in this study cooperative learning groups) improve CT (Abrami et al., 2015; Johannessen, 2001). Cooperative learning group stimulates students' thinking and provides scaffolds for them to think of new ideas and the more they engaged in the group discussions the more benefit they gain from it. This analysis also confirms the principle of integration in Merrill's First Principle of Instruction. Merrill (2002) notes in this principle that, learning takes place when the learner is given the opportunity to demonstrate the new knowledge or skill they have been recently taught publicly.

With regards to H2, the results also revealed that the systematic design of cooperative learning activities based on the principles of the cognitive test items had a significant effect on sub-CT skills. In Q3 of the interview, 5 out of the 6 students admitted that group discussion affected the way they think. They all acknowledged that they now tend to look at a problem from different perspectives. As an illustration, S2 noted that:

> *After the group study, I think my way of thinking has changed. Studying alone may not be a comprehensive way to think about a problem. After studying in a group, I think I will consciously think about a problem from many aspects.*

This was seconded by S5 and S6 who confirmed that some change in thinking has occurred in their thinking. S5 mentioned that:

> *"I used to think independently and get personal answers, but now I have learned to exchange ideas with others to find the best answer" (S5).*

Similarly, S6 said:

> *"Some changes have been made to me, such as the multi-level of thinking about a problem."* (S6)

This is a good sign for CT disposition. This is in line with the views of some key experts in the domain of CT (Ennis, 1996; Halpern, 2014) who argue that good critical thinkers are willing to change their perspective once they find out that contrary opinion is proven to be right and theirs wrong. However, not all students found cooperative learning groups affected their thinking. S4 noted that:

> *There is no change basically because in the past, I often used the method of cooperative learning groups in my study, which was carried out based on personal reflections first. In the end, it was usually just a summary of one more idea, and my way of thinking has not changed in essence.(S4)*

From students' responses in the first three questions, it is arguable that hypothesis testing, argument analysis, and problem-solving enjoyed higher mean gains after testing with a general cognitive CT assessment test at post-test due to engagement in MET problems in cooperative learning groups. The high performance in the skill of problem-solving is contrary to Dwyer et al. (2011) study in which the problem-solving skill measured by the HCTA test did not show any significant improvement in the experimental group.

Interestingly, students affirmed that the use of the *Rationale* CSAV tool also had a considerable effect on their thinking skills and disposition. In the interview Q4-Q6, some students noted that the use of the CSAV tool affected their thinking process while others thought it affected only the way they organized their ideas. 5 of the students (S1, S3, S4, S5 and S6) reported that it helped them to visualize their thinking. In order words it helped them to explicitly express their thoughts. Thus, endorsing studies that contend for this position (McGuinness, 2000; Swartz & Mcguinness, 2014). S6 for example admitted by saying:

> *"Yes, Rationale was helpful because it provides a clear flow chart of thinking, reduces chances to forget, make errors, and allows for more detailed and careful thinking."*

This trend was similar in the responses of the rest of the 4 students. However, *S2* mentioned that CSAV did not make much difference in his thinking

> *"Rationale did not help me in thinking but it helps me to organize and clarify some logical thinking".*

In both cases, CSAV was useful in one CT aspect or the other. This is in line with current literature (Hoffmann, 2018; Rider & Thomason, 2014) which support the view that CSAV promotes students' self-correction given that students have the tendency not to proofread their work. Hoffman (2018) for instance emphasized that CSAV enables and motivates students to identify weaknesses, gaps, biases, and limiting perspectives in their reasoning and to correct them. Nevertheless students found the experience of having to practice reasoning through argument visualization as a new experience. They admitted that it was a little difficult to operate the CSAV tool at the beginning and even preferred to draw by hand, but later found that it was more convenient. S2 for instance reported that:

> *I think this experience is quite new, and I was not used to reflecting on problems through argument visualization at the beginning, but after many attempts, I gradually got used to this way of thinking, and I will think about problems from different perspectives.*

S1 added that:

> *"At the beginning I found it more convenient to draw by hand. After a few times of use, after getting used to it, I felt it made the structure of my argument quite clear and coherent compared with mapping by hand, it looks better".*

And out of the five CT sub-skills, S1 commented that:

> *"I think it helped me to carry out exploration and practice in verbal reasoning, hypothesis testing, and argument analysis".*

S3 added that:

> *"CT with mapping enables looking at a problem from different perspectives as he commented "When you look at a problem, you don't just look at it superficially, you start to look critically".*

S4 complained that:

> *"It is messy at first, it takes some time to get used to it, but it's a good experience. It is helpful to learn critical thinking ability using argument visualization".*

From the analysis above, students generally agreed that in the process of learning critical thinking, they learned to think from different perspectives when discussing in their groups with other team members while the CSAV tool helped them express their thinking more clearly.

The low performance in verbal reasoning and likelihood and uncertainty could be because skills were not sufficiently explained to students in the intervention, as they were over-showed by the MET contents, which in itself had a lot of hands-on the task. It is worthy to note that the intervention took effect in the second half of the semester when the previous chapters had been covered. The intervention focused mainly on the last topics, only 40 % of the coursework was covered lasting seven sessions, out of which five were used for the teaching intervention and two sessions used for pre-and-post testing. As a result, there were not many exercises on verbal reasoning which placed the two groups on an equal level in verbal reasoning. Given that both groups started the course at the same time, and had both covered the 60 % of the course work in the first half of the semester, it is reasonable to suggest that both groups were provided with equal opportunities to engage with verbal reasoning and argument analysis in solving MET-related problems. Students in universities are expected to develop relevant CT skills as part of thier regular learning activities in the various courses they take. We are not surprised that students in the comparison group also showed some score improvment at the posttest. This interpretation is consistent with studies which found that engagement with university experience benefits CT related skills like verbal reasoning (Dwyer, Hogan, & Stewart, 2015; Lewis, Pascarella, & Terenzini, 1992). This could be the reason why the students in the CSAVIN group did not show much more greater improvement compared to the No-CSAVIN group. The duration and the number of topics covered and the quantity of exercise might not have been sufficient to enable transfer to a general domain CT test. The results may contribute to the available research evidence, which suggests that to acquire transferable CT skills can mainly be achieved via interventions that last for a prolonged period, and should cover more topics (Halpern, 2014). Future research could consider a whole semester course with the intensive and deliberate practice of CT skills by evenly using examples from other domains outside subject-matter instruction alongside subject matter contents.

Another probable cause for the absence of transfer of subject-CT skills to some of the general CT skills could be explained by the views of subject specificity advocates who argue that CT is thinking in a specific domain (McPeck, 1990). It could be contended that this study is consistent with this view. Nonetheless, the generalists' view may also argue that the main focus was not CT as a separate thread of the MET course (Ennis, 1989).

Also, the HCTA test focuses more on scenarios that are common in western cultural society. In contrast, this study was conducted in an eastern cultural society. Therefore, the participants might have limited knowledge about contents and reasoning used to formulate test items. This analysis is consistent with Stapleton (2001), who states that familiarity with the content or subject enhances CT achievement. Thus, the students or participants in question have to be sufficiently familiar with such scenarios in their everyday communities for significant results to be achieved in generalist specialists CT tests.

In terms of H3, as predicted, the analysis of students' performance in individual skills revealed that the CSAVIN group students performed significantly well in argument analysis skills in the HCTA test. The significant performance can be explained by the CSAV tool, which is mainly an argument visualization tool (Van Gelder, 2007). Students' testimonies about the benefits they got from using the CSAV tool also explains their outstanding performance in the argument analysis skill from pretesting to post-testing measured by the HCTA test. The CSAVIN group performed significantly better at posttest in the argument analysis skill because of the deliberate use of the CSAV tool. The result consistent with and contributes to the existing empirical evidence which found that the deliberate and intensive practice in CSAV enhances CT gains (Dwyer et al., 2015; Eftekhari et al., 2016; Twardy, 2004; Van Gelder, 2015). This study also emphasizes that CSAV also instills in the students the disposition to think purposefully.

7. Conclusions, limitations, and suggestions

It was assumed in this study that CSAV-based CT instructional interventions lacked some explicitness in terms of instructional design. This study designed and examined the effects of a systematic and model-based approach of designing CSAV-Infusion on overall CT skills and five sub-skills measured by the HCTA test to fill this gap. The compared results obtained from the HCTA test showed that the experimental group significantly outperformed the comparative group in overall CT ability, and especially in hypothesis testing, problem-solving, and argument analysis on the post-test than compared to the pre-test. This result suggests that combining various explicit instructional strategies in the CSAV-Infusion approach is effective in undergraduate CT development. Training a course instructor in anchored instruction, having students practice CT skills in cooperative learning groups, and using the CSAV tool to build dispositional skills towards CT in a systematic and model-based design is a promising approach to designing CSAV-based CT instruction. It could be particularly useful in a context like China, where CT instruction is not so popular in universties.

However, this study had some limitations. First of all,we acknowledge that focusing on 40 % of a single course to provide CSAV-based CT instruction for seven sessions may have some significant effect on overall CT and sub-skills assessed by a domain-general CT test. Still, given that the duration and scope of the intervention was restricted, the impact on transfer may not be significant. The transfer of CT skills requires long term training (Halpern, 2014). However, we are hoping to show to readers in this study that a relatively short duration of well-designed intervention could result in a significant growth in CT skills. We are simply indicating in this study the tendency for the students in the experimental group to improve in such short duration of intervention. If all the courses are designed the way we attempted to do and the duration of the intervention is at least for all the chapters in a course, we expect the development of CT skills to be even significantly higher than a "regular" method of teaching the university courses.

Secondly the results obtained in this study show that a combination of many explicit strategies may maximize CT improvement in a single infusion lesson, but it may also be a hindrance because students are saturated with a variety of strategies. Consequently, there is not sufficient time to reflect. Thinking tends to be rushed, which is not helpful for CT development. CT should be a reflective activity where the brain deliberately slows down to reflect on what to do or decide using those strategies that produce the desired outcomes. To achieve better results, we suggest that two or three skills may be selected to be practiced in one lesson.

Albeit, the CSAVIN group was compared with a comparison group; it was not compared to any other CT training condition, unlike Dwyer et al. (2015). Therefore, it is suggested here that future comparison studies may manipulate other control conditions to compare with a systematic and model-based design. For example, another control condition may infuse only CSAV without other strategies. At the same time, a systematic and model-based design intervention that combines any other selected strategies, and the two compared to judge which condition yields better CT performance.

Given that the Infusion approach uses several explicit strategies in subject matter CT instruction, it is arguably more cognitively demanding as the various strategies interact to achieve improvement on the targeted CT skills. The implication for the use of CSAV in future studies is that it can only be used as an additional strategy among the other numerous explicit strategies in subject-matter instruction. This contribution strongly suggests that a systematic and model-based design could be one possible way that CSAV-based CT instruction, research can address some of the deficiencies of other research methods in investigating the role of CSAV-based CT instruction.

As educational technologies continue to flood the domain of education, instructors need to be assisted, especially since those needed for CT development like CSAV are so time-consuming. The systematic and model-based design in this study provides a practical guide on how to engage an interdisciplinary team of researchers in the systematic design of learning activities geared towards CT enhancement. Future research on CT instruction in domain-specific courses like the MET course could also consider designing a subject-specific CT test to ensure the effective assessment of domain-specific CT skills.

Supplementary materials

The authors are not authorized to share some of the material and tools used, given that they are licensed. However, the *Rationale* software and platform licensed by the *Austhink* Company can be accessed through this link https://www.rationaleonline.com/, info@kritischdenken.nl, and information on Halpern Critical Thinking Assessment test could be accessed at www.schuhfried.at.

Funding

This work was supported by the Zhejiang Provincial Philosophy and Social Science interdisciplinary Key Project: "Theoretical and practical study on ICT-supported students' critical thinking skill cultivation" (19XXJC03ZD-2).

CRediT authorship contribution statement

Berty Nsolly Ngajie: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration. **Yan Li:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Funding acquisition. **Dawit Tibebu Tiruneh:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision. **Mengmeng Cheng:** Validation, Investigation, Data curation, Writing - review & editing, Project administration.

Declaration of Competing Interest

None

Acknowledgments

We would like to thank the Austhink/REASONINGLAB Company for providing us with free *Rationale* platform accounts for the students. We are also grateful to Dr. Ekwelle E. T.M, Dr. Xuesong Z., Micheal M., Rhys E., Alex D. and Kevin F. for their helpful comments and for proofreading the manuscript.

Appendix A

See Table A1

References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Thinking skills and dispositions : A stage 1 meta-analysis. *Review of Educational Research, 78*(4), 1102–1134. https://doi.org/10.3102/0034654308326084.

Table A1
Interview Questions.

Questions	Description
1	What changes did you notice in the course after the previous sessions?
2	Please tell me about your experiences of thinking about thinking in this course and whether or not it was a slow, fast or gradual process. Do you think you learned better in a group than alone?
3	After studying in groups do you feel there is a change in the way you think about thinking?
4	Did you find the *Rationale* CSAV tool affected your ability to think critically?
5	Please tell me about your overall learning experience and try to describe the nature of this experience, particularly with regards to learning about critical thinking and argument mapping.
6	Do you think that the exercises helped you acquire the critical thinking skills of verbal reasoning, thinking as hypothesis testing, argument analysis, Likelihood and uncertainty analysis, and problem-solving and decision making to complete a task?
7	Did you find any challenges with the Rationale argument mapping platform?

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research, 85*(2), 275–314. https://doi.org/10.3102/0034654314551063.

Alan Bensley, D., & Spero, R. A. (2014). Improving critical thinking skills and metacognitive monitoring through direct infusion. *Thinking Skills and Creativity, 12*, 55–68. https://doi.org/10.1016/j.tsc.2014.02.001.

Angeli, C., & Valanides, N. (2009). Instructional effects on critical thinking: Performance on ill-defined issues. *Learning and Instruction, 19*(4), 322–334. https://doi.org/10.1016/j.learninstruc.2008.06.010.

Bailin, S. (2002). Critical thinking and science education. *Science Education, 11*, 361–375. https://doi.org/10.1023/A:1016042608621.

Bensley, D. A. (2010). A brief guide for teaching and assessing critical thinking in Psychology. *Psychological Science, 23*(10), 1–17.

Bezanilla, M. J., Fernández-nogueira, D., Poblete, M., & Galindo-domínguez, H. (2019). Methodologies for teaching-learning critical thinking in higher education : The teacher’ s view. *Thinking Skills and Creativity, 33*(July), 100584. https://doi.org/10.1016/j.tsc.2019.100584.

Chen, Y., & Yu, L. (2018). Research on critical thinking and cultivation strategy of college students in China. December *2nd International Conference on Advanced Education and Management Science (AEMS 2018)*. https://doi.org/10.12783/dtssehs/aems2018/28054.

Clark, R. C., & Mayer, R. E. (2011). *E-Learning and the science of instruction* (3rd ed.). San Francisco, CA: Wiley & Sons, Inc.

Davies, M. (2011). Concept mapping, mind mapping and argument mapping: What are the differences and do they matter? *Higher Education, 62*(3), 279–301. https://doi.org/10.1007/s10734-010-9387-6.

Davies, M. (2012). Computer-Aided Mapping and the teaching of critical thinking. *Inquiry Critical Thinking Across the Disciplines, 27*(2), 15–30. https://doi.org/10.5840/inquiryct20122729.

Davies, M. (2013). Critical thinking and the disciplines reconsidered. *Higher Education Research & Development, 32*(4), 529–544. https://doi.org/10.1080/07294360.2012.697878.

Davies, W. M. (2006). An’ infusion’ approach to critical thinking: Moore on the critical thinking debate. *Higher Education Research & Development, 25*(2), 179–193. https://doi.org/10.1080/07294360600610420.

Davies, W. M. (2009). Computer-assisted argument mapping: A rationale approach. *Higher Education, 58*(6), 799–820. https://doi.org/10.1007/s10734-009-9226-9.

Davies, M., Barnett, A., & Gelder, T. V. (2019). Using computer-aided argument mapping to teach reasoning. In J. A. Blair (Ed.), *Studies in critical thinking* (pp. 131–176). Retrieved from https://philpapers.org/rec/DAVUCA-2.

Dewey, J., & Bento, J. (2009). Activating children’s thinking skills (ACTS): The effects of an infusion approach to teaching thinking in primary schools. *The British Journal of Educational Psychology, 79*(2), 329–351. https://doi.org/10.1348/000709908X344754.

Dwyer, C., Hogan, M., & Stewart, I. (2011). The promotion of critical thinking skills through argument mapping. In C. P. H, & J. M. Forte (Eds.), *Critical thinking* (pp. 97–121). Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84861344240&partnerID=tZOtx3y1.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2012). An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments. *Metacognition and Learning, 7*(3), 219–244. https://doi.org/10.1007/s11409-012-9092-1.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2013). An examination of the effects of argument mapping on students’ memory and comprehension performance. *Thinking Skills and Creativity, 8*, 11–24. https://doi.org/10.1016/j.tsc.2012.12.002.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2015). The effects of argument mapping-infused critical thinking instruction on reflective judgement performance. *Thinking Skills and Creativity, 16*, 11–26. https://doi.org/10.1016/j.tsc.2014.12.002.

Eftekhari, M., Sotoudehnama, E., & Marandi, S. S. (2016). Computer-aided argument mapping in an EFL setting: does technology precede traditional paper and pencil approach in developing critical thinking? *Educational Technology Research and Development, 64*(2), 339–357. https://doi.org/10.1007/s11423-016-9431-z.

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher, 18*(3), 4–10. https://doi.org/10.3102/0013189x018003004.

Ennis, R. H. (1990). The extent to which critical thinking is subject-specific: Further clarification. *Educational Researcher, 19*(4), 13–16. https://doi.org/10.3102/0013189X019004013.

Ennis, R. H. (1996). Critical thinking dispositions : Their nature and assessability. *Informal Logic, 18*(2&3), 165–182.

Facione, P. A. (2020). Critical thinking: What it is and why it counts. *Insight Assessment, 53*(9). https://doi.org/ISBN 13: 978-1-891557-07-1.

Fischer, F., Kollar, N., Ufer, S., Sodian, B., Heinrich Hussmann, R. P., Neuhaus, B., et al. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in Education. *Frontline Learning Research, 5*, 28–45. https://doi.org/10.14786/flr.v2i2.96.

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). https://doi.org/10.1017/CBO9781107415324.004.

Harrell, M., & Wetzel, D. (2015). Using argument diagramming to teach critical thinking in a first-year writing course. In M. Davies, & R. Barnett (Eds.), *The palgrave handbook of critical thinking in Higher education* (pp. 1–25). https://doi.org/10.1057/9781137378057.

Hitchcock, D. (2017). The effectiveness of instruction in critical thinking. In D. Hitchcock (Ed.), *On reasoning and argument. argumentation library* (pp. 499–510). https://doi.org/10.1007/978-3-319-53562-3_33.

Hoffmann, M. H. G. (2018). Stimulating reflection and self-correcting reasoning through argument mapping: Three approaches. *Topoi, 37*(1), 185–199. https://doi.org/10.1007/s11245-016-9408-x.

Huan, L., & Hu, Y. (2014). The mapping knowledge domains based on the hotspots research of critical thinking in China. *Advances in Psychology, 04*(06), 857–862. https://doi.org/10.12677/ap.2014.46111.

Johannessen, L. R. (2001). Teaching thinking and writing for a new century. *The English Journal, 90*, 38–46.

Koc, E. W., Kahn, J., Koncz, A. J., Salvadge, A., & Longenberger, A. (2019). *Job outlook 2019*. Retrieved from https://www.odu.edu/content/dam/odu/offices/cmc/docs/nace/2019-nace-job-outlook-survey.pdf.

Ku, K. Y. L. (2009). Assessing students’ critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity, 4*(1), 70–76. https://doi.org/10.1016/j.tsc.2009.02.001.

Lewis, L. S., Pascarella, E. T., & Terenzini, P. T. (1992). How college affects students: Findings and insights from twenty years of research. *Academe, 78*(4), 44. https://doi.org/10.2307/40250363.

Lo, C. K., & Hew, K. F. (2017). Using’ first principles of instruction’ to design secondary school mathematics flipped classroom: The findings of two exploratory studies. *Educational Technology & Society, 20*(1), 222–236.

Marin, L. M., & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity, 6*(1), 1–13. https://doi.org/10.1016/j.tsc.2010.08.002.

Marsden, E., & Torgerson, C. J. (2012). Single group, pre- and post-test research designs: Some methodological concerns. *Oxford Review of Education, 38*(5), 583–616. https://doi.org/10.1080/03054985.2012.731208.

McGuinness, C. (2000). ACTS A methodology for enhancing thinking skills across-the-curriculum. *Teaching Thinking*, (5), 1–16. Retrieved from http://www.sustainablethinkingclassrooms.qub.ac.uk/ACTS_for_Teach_Think.pdf.

McPeck, J. E. (1990). Critical thinking and subject specificity: A reply to Ennis. *Educational Researcher, 19*(4), 10–12. https://doi.org/10.3102/0013189x019004010.

Mercer, N., Hennessy, S., & Warwick, P. (2019). Dialogue, thinking together and digital technology in the classroom: Some educational implications of a continuing line of inquiry. *International Journal of Educational Research, 97*, 187–199. https://doi.org/10.1016/j.ijer.2017.08.007.

Merrill, M. D. (2002). First principles of instruction. *ETR&D, 50*(3), 43–59. https://doi.org/10.1007/BF02505024.

Moore, T. (2004). The critical thinking debate : how general are general thinking skills? *Higher Education Research & Development, 23*(1), 3–18. https://doi.org/10.1080/0729436032000168469.

Rapanta, C., & Walton, D. (2016). The use of argument maps as an assessment tool in higher education. *International Journal of Educational Research, 79*, 211–221. https://doi.org/10.1016/j.ijer.2016.03.002.

Rider, Y., Thomason, N., et al. (2014). Cognitive and pedagogical benefits of argument mapping. In A. Okada (Ed.), *Knowledge Cartography: Software Tools and Mapping Techniques, Advanced Information and Knowledge Processing* (pp. 113–134). London: Springer-Verlag. https://doi.org/10.1007/978-1-84800-149-7.

Stapleton, P. (2001). Assessing critical thinking in the writing of Japanese university students. *Written Communication, 18*(4), 506–548. https://doi.org/10.1177/0741088301018004004.

Sun, M., Wang, M., & Wegerif, R. (2019). Using computer-based cognitive mapping to improve students' divergent thinking for creativity development. *British Journal of Educational Technology, 50*(5), 2217–2233. https://doi.org/10.1111/bjet.12825.

Swartz, R., & Mcguinness, C. (2014). *Developing and assessing thinking skills, the International Baccalaureate Project 2014 final report, part 1: Literature review and evaluation framework*. https://doi.org/10.13140/RG.2.1.4917.6163.

Tian, J., & Low, G. D. (2011). Critical thinking and Chinese university students: A review of the evidence. *Language Culture and Curriculum, 24*(1), 61–76. https://doi.org/10.1080/07908318.2010.546400.

Tian, J., & Low, G. D. (2012). To what extent are postgraduate students from China prepared for academic writing needed on UK master's courses? *Language Culture and Curriculum, 25*(3), 299–319. https://doi.org/10.1080/07908318.2012.734313.

Tiruneh, D. T., Verburgh, A., & Elen, J. (2014). Effectiveness of critical thinking instruction in higher education: A systematic review of intervention studies. *Higher Education Studies, 4*(1), 9–44. https://doi.org/10.5539/hes.v4n1p1.

Tiruneh, D. T., Mieke, D. C., Spector, J. M., Gu, X., & Elen, J. (2017). Toward a systematic and model-based approach to design learning environments for critical thinking. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, design, and technology*. https://doi.org/10.1007/978-3-319-17727-4.

Tiruneh, D. T., De Cock, M., & Elen, J. (2018). Designing learning environments for critical thinking: Examining effective instructional approaches. *International Journal of Science and Mathematics Education, 16*(6), 1065–1089. https://doi.org/10.1007/s10763-017-9829-z.

Twardy, C. R. (2004). Argument maps improve critical thinking. *Teaching Philosophy* (Vol. 27,, 95–116. https://doi.org/10.5840/teachphil200427213.

Van Gelder, T. (2007). The rationale for rationale™. *Law Probability and Risk, 6*(1–4), 23–42. https://doi.org/10.1093/lpr/mgm032.

Van Gelder, T. (2015). Using argument mapping to improve critical thinking skills. In M. Davies, & R. Barnett (Eds.), *The palgrave handbook of critical thinking in Higher education* (pp. 183–192). https://doi.org/10.1007/978-1-137-37805-7_12.

Yanhong, R., & Lin, T. (2014). The critical thinking and Chinese creative education. *Canadian Social Science, 10*(6), 206–211. https://doi.org/10.3968/4552.

Zhang, T. (2017). Why do Chinese postgraduates struggle with critical thinking? Some clues from the higher education curriculum in China: Some clues from the higher education curriculum in China. *Journal of Further and Higher Education, 9486*, 1–15. https://doi.org/10.1080/0309877X.2016.1206857.