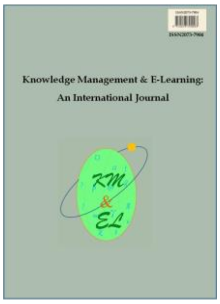
Knowledge Management & E-Learning, Vol.14, No.2. Jun 2022

Identifying critical thinking skills used by experts versus novices to construct argument maps in a computer-aided mapping tool

Allan Jeong
The Florida State University, USA
Hae Young Kim
The Florida State University, USA



Knowledge Management & E-Learning: An International Journal (KM&EL)
ISSN 2073-7904

Recommended citation:
Jeong, A., & Kim, H. Y. (2022). Identifying critical thinking skills used by experts versus novices to construct argument maps in a computer-aided mapping tool. *Knowledge Management & E-Learning*, *14*(2), 125–149. https://doi.org/10.34105/j.kmel.2022.14.008

Knowledge Management & E-Learning, 14(2), 125–149

Identifying critical thinking skills used by experts versus novices to construct argument maps in a computer-aided mapping tool

Allan Jeong*

Faculty of Education
The Florida State University, USA
E-mail: ajeong@fsu.edu

Hae Young Kim

Faculty of Education
The Florida State University, USA
E-mail: hkim7@fsu.edu

*Corresponding author

Abstract: Research shows that using computer-aided mapping tools improves critical thinking skills, but prior research provides limited evidence to show how the use of specific critical thinking skills increases map quality. This qualitative study observed 4 experts and 5 novices use a computer-aided mapping tool to construct argument maps. The analysis of video recordings with think-aloud protocols and retrospective interviews revealed the use of a five-step argument mapping process (read claims, position conclusion, position claims, link claims, revise links) with the experts using a more sequential application of the five-step process and producing more accurate maps than novices. The novices showed the tendency to position and link claims as a joint action, making map revision more cumbersome. The experts exhibited the tendency to work backward from conclusion to claim while the novices exhibited the reverse tendency. This study's findings identify processes that differentiate experts from novices and validate specific thinking skills that can be used to improve map quality, and how these processes can be operationalized in terms of discrete mapping behaviors performed on screen that can be mined and analyzed in mapping tools to assess and diagnose students' mapping skills.

Keywords: Critical thinking; Concept mapping; Argument analysis

Biographical notes: Dr. Allan Jeong is an Associate Professor in the Department of Educational Psychology and Learning Systems at Florida State University, developing learning analytic tools and methods to mine, assess, and model the individual and collaborative processes of critical thinking and problem solving.

Dr. Hae Young Kim is an Adjunct Instructor in the Faculty of Education at Florida State University. She works as a psychometrician and an instructional designer at a non-profit organization in Florida.

1. Introduction

Critical thinking is an essential skill in higher education and professional work, but research shows that many college students fail to develop critical thinking skills and effectively use the skills (Davies, 2011). A large proportion is not able to distinguish fact from opinion, objectively analyze two conflicting viewpoints, and analyze problems to determine underlying causes without influence from appeals to emotion (Roksa & Arum, 2011), distinguish claims from reasons stated to support a claim, cannot select correct reasons to support their claims (Larson, Britt, & Larson, 2004), identify structural flaws in arguments, and correctly distinguishing claims that are versus are not backed by warrants (Larson, Britt, & Kurby, 2009).

To address these deficiencies, argument analysis is being used to teach critical thinking across disciplines because argumentation is essential in the scientific and problem-solving process (Cottrell, 2017; Davies, Barnett, & van Gelder, 2019). Argument analysis is a process of evaluating evidence, drawing appropriate conclusions, distinguishing arguments from non-arguments, and finding assumptions, identifying the functional roles and hierarchical relationships between propositions, and evaluating the truth-value of claims in relation to its minor claims (Toulmin, 1958). Difficulties arise when premises are not explicitly stated, requiring one to infer missing premises to establish their inter-relationships (Ennis, 1982). Altogether, argument analysis is a complex process (Weinerth, Koenig, Brunner, & Martin, 2014) that requires high levels of attention, memory, and cognitive effort (Harrell, 2007; van Bruggen, Kirschner, & Jochems, 2002).

Because argument analysis is a highly complex process and requires much cognitive effort, mapping tools have been used to help students construct maps to identify and map out the relationships between premises and claims (van den Braak, Oostendorp, Prakken, & Vreeswijk, 2006), and in the process of doing this, help students develop a deeper understanding of learned concepts in terms of how they are applied to solving a program (Wu, Wang, Kirschner, & Spector, 2018). Tools like Argument Mapper (Wright, Sheffield, & Santosa, 2017), AVIZE (Green, Branon, & Roosje, 2019) and Rationale (van Gelder, 2007) enable students to position and link nodes to visually map out complex hierarchical relationships between premises and claims, while dual mapping environments (Chen, Wang, Dede, & Grotzer, 2021; Wu & Wang, 2012) enable students to create and link concepts in an adjacent concept map to specific premises, claims, tasks, or events presented in an argument or "reasoning" map. Analyzing complex arguments with diagrams reduces cognitive load (Correia & Aguiar, 2014) and allocates more working memory to interpret text, identify functional elements (claims, supports, objections, counterarguments, etc.), and test hierarchical relationships (Harrell, 2007; van Bruggen, Kirschner, & Jochems, 2002) – all of which are higher-order cognitive skills (Aguiar & Correia, 2017). Furthermore, tools like REASON (ThinkReliability, 2007) prescribe the use of backward as a goal-driven approach to diagramming arguments (Sharma, Tiwari, & Kelkar, 2012) whereas Betty Brain steps students through the breadth-first process to help students create better maps (Biswas, Segedy, & Bunchongchit, 2016).

In general, studies show that using mapping tools improves learning with moderate to large effect sizes (Schroeder, Nesbit, Anguiano, & Adesope, 2017) and improves students' critical thinking skills (Eftekhari, Sotoudehnama, & Marandi, 2016; van Gelder, 2015). Students must apply more complex cognitive processes (combined comprehension, construction and interpretation) when constructing maps as opposed to simply studying a given map (Easterday, Aleven, & Scheines, 2007). This explains in part why college students' use of argument mapping software has been found to improve their critical thinking skills (van Bruggen, Boshuizen, & Kirschner, 2003; Harrell, 2011; Twardy, 2004; van Gelder, 2007). Even so, students often feel overwhelmed and lose motivation while constructing maps (Beitz, 1998; Kinchin, 2001). To address this problem, introductions to mapping tools are often accompanied with instruction on specific thinking strategies to manage the complexity of the tasks. These strategies include directing students to place the goal at the top (Eppler, 2006), sorting before linking nodes (Aguiar & Correia, 2017), sorting nodes by level of generality (Cañas, Reiska, & Möllits, 2017), positioning nodes with reading flow (Aguiar & Correia, 2017; Jeong & Lee, 2012), using five whys, backward chaining and depth-first method (Al-Ajlan, 2015; Chen, Li, & Shady, 2010), and using a breadth-first process to review maps (Biswas, Segedy, & Bunchongchit, 2016).

However, prior research has yet to fully identify and measure the extent students use particular strategies while constructing maps (Wang, 2019) and to what extent does the use of specific strategies produce higher quality maps (Schroeder, Nesbit, Anguiano, & Adesope, 2017). Prior research focus primarily on evaluating the effects of using computer-based mapping tools on map quality (Cañas, Novak, & Reiska, 2015) and students' higher-order skills (Cañas, Reiska, & Möllits, 2017) without measuring and identifying the precise strategies students used to construct their maps. Among the studies that compare mapping tool techniques and interventions, these studies do not isolate the specific components within the intervention nor map the components to specific thinking processes to determine the effects of the thinking process on map quality. One study compared thinking processes used by experts versus novices to construct causal diagrams, but this study focused only on identify errors in reasoning as students interpreted their diagrams (Easterday et al., 2009). This gap in our current understanding of the link between thinking processes and map quality help explain the high variance consistently observed in map quality regardless of the efficacy of an instructional intervention (van den Braak et al., 2006). This research gap can be blamed in part on the difficulties of operationally defining and measuring specific processes used to construct maps and the difficulties in assessing the finer-grain qualities of maps associated with the use of specific skills or strategies. As a result, research is needed to validate and establish the relationship between the use of specific skills and map quality (Kuhn & Udell, 2003).

To address this problem, this qualitative study identifies and compares the processes that participants with high versus low prior knowledge of argument analysis used to construct argument maps and examines how the observed processes are linked to map quality. At the same time, this study identifies which and to what extent the thinking processes and strategies used by the participants (internal mental processes) are tied to discrete map construction behaviors (external and observable processes) that can be mined and analyzed in computer-aided mapping tools to automate the identification, verification, and assessment of the reasoning skills learned and used by students to construct their maps.

1.1. Research on reasoning strategies and constructing maps

At this time, empirical research on reasoning and map construction strategies and their effects on map quality is very limited (Cañas, Reiska, & Möllits, 2017; Schroeder et al., 2017). Some of the strategies that can be used to construct maps are strategies used in problem solving, or more specifically, strategies for constructing explanations on how ideas or events are logically or causally connected. One such strategy is the depth-first process when linking A to B and linking B to C are performed back-to-back (or linking B to C and then linking A to B) to produce the chain A--> B --> C. This process is best suited for diagnosing and identifying root causes to problems (Chen et al., 2010; Serrat, 2017; Terjesen & Patel, 2017). In contrast, the breadth-first process takes place when linking A to C and linking B to C are performed back-to-back to identify two claims that directly support C. This strategy is prescribed as the process of choice when working top-down from a given goal (Sharma, Tiwari, & Kelkar, 2012). Because depth-first places less demand on memory than breadth-first based on machine learning research (Al-Ajlan, 2015), depth-first and backward processing can be used in conjunction when given a clear outcome (Sharma, Tiwari, & Kelkar, 2012).

The depth-first strategy can be implemented by using backward processing, when analysis starts with the outcome or conclusion, then works backward to find supporting facts or claims (Sharma, Tiwari, & Kelkar, 2012) - a process also useful for diagnosing tasks (Al-Ajlan, 2015). The depth-first strategy can also be implemented by using forward processing as a data-driven approach for predicting possible outcomes through the iterative application of logic rules to given facts (Hinkelmann, 2004) - a process well suited for analyzing tasks with no distinct goals such as planning, designing, and monitoring process (Sharma, Tiwari, & Kelkar, 2012). When faced with unfamiliar problems, people tend to use forward processing by default to evaluate arguments (Heit, 2007; Oaksford & Hahn, 2007) because making predictions is a process that can instill a higher sense of confidence and judgment (Sweller, Clark, & Kirschner, 2011; Tversky & Kahneman, 1980).

At this time, three studies are known to have operationalized, measured, and examined to what extent students' use of these thinking strategies correlates with map quality. One qualitative study revealed that the group producing the most accurate causal map used more backward than forward processing (Lee, 2012). Less accurate maps produced by the other two groups exhibited no clear tendency to use one process over the other. Shin and Jeong's (2021) qualitative analysis found that the students that constructed better causal maps (when given a specific outcome) showed the tendency to use more backward than forward processing (high backward/forward ratio) and more breadth-first than depth-first processing (high breadth/depth ratio). Students' prior knowledge of the content of the causal maps was not found to be associated with which processes they used. Finally, Jeong (2020) analyzed the movements and placements of nodes captured in trace data recorded by a computer-aided mapping tool to also find that students who use a higher ratio of backward to forward processing, and a higher ratio of using depth-first to breadth-first processing constructed better maps.

These preliminary studies, however, provide only correlational data to illustrate how students' use of specific strategies help them to construct higher quality maps. The data used in the Jeong (2020) study to infer what strategies students were using was limited to and based only on the computer-based mapping actions (e.g., place node immediately below vs to the right of the previously moved node) students performed while constructing their maps. No verbal protocol data was collected to validate the inferred link between a specific mapping action and the use of a specific strategy. Furthermore, what strategies that are used to construct maps may be influenced by one's prior knowledge of maps given that Körner (2005) found that students with low prior knowledge of hierarchical maps were less able to comprehend maps. To fully identify and validate the strategies that are being used to improve map quality, the aim of this study was to identify, measure, and compare the extent to which specific strategies are used to construct argument maps between experts and novices with high versus low prior knowledge with argument analysis and diagrams.

1.2. Research questions

The purpose of this qualitative study was to identify and compare the processes and strategies participants with high versus low prior knowledge of argument analysis used to construct argument maps and to examine which processes are used to produce higher quality maps. The second purpose of this study was to identify which and to what extent the observed strategies (internal processes) are tied to specific mapping behaviors that are external, observable, and can be mined in mapping tools to automate the identification and assessment of students' reasoning skills. As a result, the research questions addressed in this study were the following:

1. What thinking processes are used by experts and novices with and without prior experience with argument analysis, respectively, to construct an argument map?
2. How do the processes and strategies used by experts to create higher quality maps differ from those used by the novices?
3. To what extent do the discrete mapping actions performed on screen denote specific cognitive acts and thinking processes used to construct argument maps?

2. Method

2.1. Participants

The participants were recruited from a large university in the U.S. southeast region. Based on Nielsen's (1994) recommendation that a sample size of 4 ±1 is sufficient for exploratory studies using think-aloud protocols, five experts and five novices were recruited for this study. A survey invitation was distributed via email and class visits with faculty and graduate students across multiple departments across campus to recruit participants. The survey requested the following information: age, gender, student or instructor, instructor's field of expertise, number of years teaching, and prior knowledge of six established e-learning design principles (Clark, 2002) that served as the main content of the argument mapping task, and prior experience using argument mapping software. Using a purposeful and convenience sampling from the pool of volunteers, the five experts were selected on the basis that they teach or have taught courses on argumentation and/or have had prior training on argument analysis. The novices were selected from volunteers who were graduate students that had little or no formal training in argument analysis.

All five experts (5 males; 42, 32, 52, 54, and 29 years in age) teach courses on argumentation with argument diagrams and/or formal reasoning (one from Criminal Justice, four in Philosophy) with 8, 24, 2, 25, and 2 years of teaching experience. Four of the experts taught courses on argumentation (one taught only formal reasoning), three had prior experience mapping arguments, none had prior experience using computer-based mapping tools, and none had prior knowledge of the e-learning design principles. The five novices (1 male, 4 females; 24, 26, 24, 24, and 54 years in age) were graduate students in Career Counseling, Mental Health Counseling, Information Science, Information Science, and Education. One novice had a prior course on argumentation, two had familiarity with 1 and 2 of the e-learning principles, and none had prior experience with mapping arguments and argument mapping tools.

2.2. Argument mapping task

Each participant was presented a 2-minute video on how to use the computer-aided mapping tool, jMAP, to move nodes, insert and delete arrows to link nodes, re-route arrows to point to a different node, and change arrow color (black for supporting premise, red for opposing premise). Next, the participants completed a 10-minute practice exercise to construct a mini argument map consisting of five nodes (claims about the importance of critical thinking in college students) while talking aloud to familiarize themselves with the talk-aloud protocol. With the auto-link feature disabled in jMAP, participants did not have to link nodes when moved into position - making the nodes easier to reposition, sort, and re-sort nodes in and out of groups while exploring their inter-relationships. This process in particular, when used to complete jigsaw puzzles, has been found to facilitate exploration and simplifies the problem space to reduce cognitive load and improve performance (Antle, 2013).

Next, participants were presented with a printed copy of a 1017-word summary description of six e-learning principles (including graphic illustrations) and their impact on learning from an article extracted from Clark (2002). The average reading times were 5 minutes 9 seconds and 4 minutes 25 seconds for the experts and novices, respectively. Participants were then presented with jMAP populated with 14 claims and one conclusion (Fig. 1) drawn from the presented summary with one claim added as a distractor item. Participants were instructed to construct an argument map to reveal the logical structure between the claims and the conclusion. Each participant talked out loud to report their thoughts while performing the task. As the researcher observed the participant perform the mapping task, the researcher reminded participants to talk out loud when needed and aided the participants only when participants experienced technical difficulties with the software. The verbal reports and actions performed on screen were recorded using a video screen-capture program and all on-screen mapping actions were logged in the jMAP application. The novices and experts spent an average of 21 minutes (min = 6:41, max = 32:29) and 26 minutes (min = 18:17, max = 34:10), respectively, to complete the argument map.

2.3. Retrospective interview session

After completing the map, each participant started the one-on-one interview session by responding to an open-ended question about the argument mapping process and experience. The participants were able to refer to their argument map on screen to help answer questions and recall the process they used to construct the map. When participants described a particular process, they were probed for more details and retrospective descriptions of the specific actions they performed (and the reasons behind the action) to identify the relationships between claims. They were also probed to explain any difficulties they experienced while trying to identify the relationships, and to share any thoughts that they may have filtered out and did not verbalize during the mapping task.

2.4. Data analysis overview

Altogether, the data collected in this study consisted of the argument maps and map scores, mapping actions logged in jMAP, video recordings of each participant's jMAP screen with think-aloud protocols, and retrospective interviews. The data from one expert (who taught formal reasoning, did not teach argumentation, and had no prior experience mapping arguments and using mapping software) was omitted because the expert exhibited high levels of frustration and discomfort, asked frequent questions about the mapping software, and required frequent assistance from the researcher during the mapping task. Using data from the remaining 4 experts and the 5 novices, the data analysis was conducted in the following sequence.

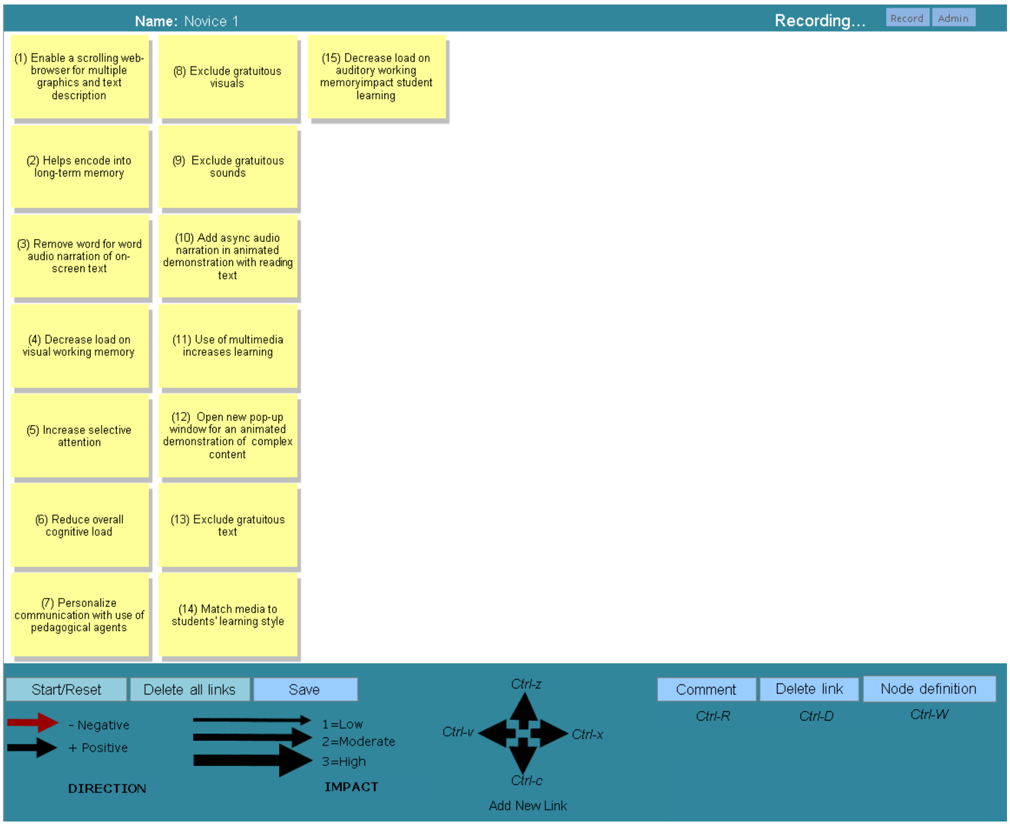


Fig. 1. jMAP screen at the start of the argument mapping task

The argument maps were evaluated and scored by the jMAP software against a criterion argument map. The retrospective interviews were then reviewed to qualitatively identify both global and atomistic processes that the participants reportedly used to construct their maps. The global processes identified from the retrospective interviews were listed in a table to reveal the componential tasks and task sequences used to complete a map. The more atomistic processes (including the map scores) were instead placed into a matrix (process X participants) to help reveal possible differences between groups.

Next, the verbal protocols and the mapping actions revealed in the screen recordings were transcribed and entered directly into each participant's jMAP log archived in a spreadsheet so that each entry was listed alongside any coinciding mapping actions (move node, location of moved node, link node, relink node, delete link, set link attribute) recorded, and time stamped in the jMAP log. The verbal protocols were parsed into units of meaning and coded within each participant's spreadsheet to identify behavioral patterns that connotate specific thinking processes. The codes were then used to locate verbal excerpts from the spreadsheet to verify and illustrate specific processes identified from the retrospective interviews. The log data was not analyzed in this study, but it will be analyzed in a separate study where algorithms will be developed to process the log data in a way that accurately identifies and measures the thinking processes identified and validated in this current study.

2.5. Scoring argument maps

The maps were scored with the jMAP software using a criterion map (Fig. 2) produced by a professor who teaches courses on multimedia design and conducts research on argumentation and assessed across five criteria (Table 1). Given that the linkages inserted into an argument map reflects the student's logical understanding between claims and premises (Chen et al., 2021), one point was awarded for correctly identifying the main conclusion or hypothesis, one point for correctly identifying each relationship between claim and hypothesis, and one point for identifying each relationship between two claims. To measure and place more weight on students' depth of analysis or the extent to which students are able to infer and articulate the logical pathways connecting root claims (or assumptions) to the main conclusion (Suthers & Hundhausen, 2003; Jeong, 2020), one point was also awarded for: a) identifying the lowest level or root claims (claims with no child claims); b) linking two chained claims branching from each root claim (R→A→B); c) linking three consecutive claims branching from the root claims(R→A→B→C); and d) linking four consecutive claims branching from the root claims (R→A→B→C→D). In this study, the links students used to connect claims were directional links identified with an arrowhead to graphically convey what would be a link label with the text "supports" or "enables" or "verifies". As a result, 0.5 point was deducted for inserting arrows pointing in the wrong direction, and 0.5 point was deducted for each arrow with an incorrect valence (positive or negative). Given that the hypothesis and claims were all provided in advance at the start of the argument mapping task, the argument diagrams were not scored on the number of claims identified in the map (Van Drie et al., 2005), the relevance and correctness of the claims (Janssen et al., 2010), and on their sufficiency and completeness (Chen et al., 2021). The map scores varied widely, ranging from 5.5 to 31 points, with all experts scoring higher than the novices.

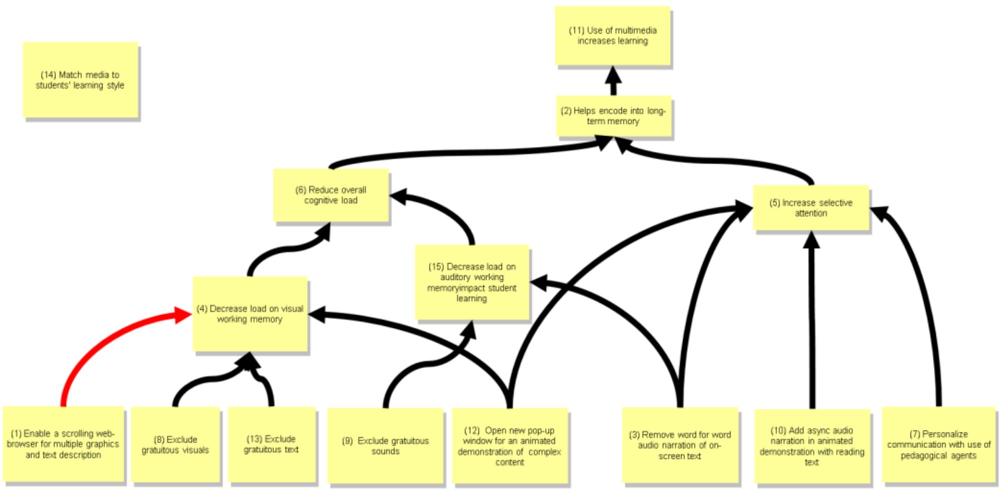


Fig. 2. Criterion map used to score participants' argument maps

Table 1
Participants' argument map scores sorted from highest to lowest score

	Correctly identified the main conclusion	# of root claims correctly identified	# of correct 1st order chain	# of correct 2nd order chain	# of correct 3rd order chain	# of correct 4th order chain	Total score
Max score	1	8	10	10	10	10	49
Expert 3	1	8	7	5	5	5	31
Expert 4	1	8	5	4	4	4	26
Expert 1	1	7	5	5	1	0	19
Expert 2	1	8	5	2	0	0	16
Novice 1	0	6	4	4	0	0	14
Novice 3	1	7	1	1	1	0	11
Novice 5	1	7 (5*)	2 (1*)	1	1	0	9
Novice 2	1	5	1	0	0	0	7
Novice 4	0	3	3 (1*)	0	0	0	5.5

Note. * Links incorrectly identified in direction and/or valence resulting in a 0.5 point deduction

2.6. Coding retrospective interviews and verbal protocols to identify processes

After reviewing the retrospective interviews and identifying the processes the participants used to construct their maps, the grounded theory approach (Guba & Lincoln, 1994) was used to identify the categories for coding the verbal protocols emerged through iterative examination of the video recordings of a randomly selected novice and expert. All codes were entered into the participant's spreadsheet alongside the current entries, mapping actions, and time stamps. New categories were added to the coding scheme when new actions could not be assigned to an existing category. Once the draft coding scheme was completed, a second coder was trained to code the videos of the selected expert and novice using the initial coding scheme. Inter-rater reliability between the second coder and the researcher was tested, disagreements were addressed, and further revisions were made to the coding scheme to resolve the disagreements. Using the revised and final coding scheme, the two coders coded the videos of another novice and another expert. The Cohen's kappa inter-rater reliability (Cohen, 1960) for coding the novice and expert data was .78 and .97, respectively, indicating substantial agreement between raters. The remaining videos were coded by the researcher using the final coding scheme.

A thematic analysis was performed on the qualitative data collected from the interviews and verbal protocols to identify verbal-cognitive actions performed during the mapping process. Thematic analysis is a method used in qualitative studies to identify meaningful patterns and themes in rich complex data in relation to the research question (Braun & Clarke, 2006). Using an inductive approach, this study started by noting verbal-cognitive actions during data analysis and continued until all exclusive actions were identified. The entire data set and coded extracts were constantly and iteratively re-examined to obtain emergent actions, generate categories of actions, and identify and relate categories to subcategories, while using the broadest categories to generate possible themes and refining the themes to integrate all coded data. After creating the initial coding scheme, a second coder was trained and applied the coding scheme to code the video recordings and protocols from one expert and one novice. The codes were compared between the researcher and second coder, and all discrepancies were discussed and resolved with revisions and improvements made to the coding scheme. The process of coding and comparing the codes on the video protocols of another expert and another novice was repeated until a high inter-rater reliability Cohen's Kappa score was achieved. Commonalities and differences noted between the codes determined the final list of verbal-cognitive actions presented in the final coding scheme.

The analysis revealed actions to show, for example, that the participant to varying degrees used breadth-first and depth-first processes, with depth-first processes performed with backward and forward chaining process as documented in prior studies (Lee, 2012, Shin & Jeong, 2021; Jeong, 2020). Identifying and counting the number of times experts and novices used the breadth-first, depth-first, backward, and forward process relied on observing the location of the moved claim relative to the location of the most recently moved claim. For example, depth-first processing occurs when A, B, and C are linked or placed in a sequence to produce the ABC chain by performing forward processing two times in succession by placing B after A, then placing C after B, or by performing backward processing two times in succession by placing B before C, then placing A before B. In contrast, breadth-first processing occurs when A is moved to the left of C using backward processing to produce A → C and then D is moved to the left of C using backward processing to produce D → C to create two stems branching from C. Breadth-first processing can also occur when C is placed to the right of A using forward processing to produce A → C and then D is placed to the immediate left of C using backward processing to produce D → C to again create two stems branching from C.

2.7. *Identifying differences between expert vs. novice processes*

Qualitative differences in the processes used by the experts and novices were determined by entering the coded behaviors (including map scores) exhibited by each expert and novice into a participant X pattern matrix. Frequency counts of specific behaviors were added to the matrix for making global comparisons that might discern possible differences in the processes used by experts versus novices. The resulting matrix provided a summary report to first reveal any notable similarities exhibited by the experts as a group and any similarities exhibited by the novices as a group. The matrix was then used to determine which similar behaviors exhibited among the experts were unique to only the experts and were not overall exhibited by the novices, and vice versa.

3. Discussion of main findings

This section starts with the presentation of the findings (research question 1) on the processes used by experts and novices to construct their argument maps, beginning with a description of global processes (or main steps in the mapping process) followed with descriptions of more local processes observed within each mapping step. Immediately following the description of specific processes are discussions of observed differences in the extent to which the specific cognitive acts and/or processes (defined by specific sequences of cognitive acts) were exhibited by the experts relative to the novices (research question 2). This discussion then moves on to the presentation of the findings (research question 3) revealing the extent to which the identified processes and cognitive acts can be denoted by discrete mapping behaviors performed on the mapping screen. Discussion of the instructional implications and directions for future research are presented to conclude the section.

3.1. The process used by experts and novices to map arguments

The analysis of both the retrospective interview and video recordings of the verbal protocols revealed that the experts and novices constructed their maps using an array of actions (some performed using specific strategies) that can be classified into five main steps: 1) scan all claims, 2) identify and place conclusion, 3) move and position claims, 4) link related claims, 5) review & correct links between claims. Table 2 summarizes the five main steps illustrated with quotations from the verbal reports and interviews and presented with some of the strategies used to perform a given step. All five of these steps mirror the 5-step process for constructing good concept maps prescribed by (Aguiar & Correia, 2017).

Table 2
Five main steps with associated strategies used to construct the argument maps

Steps	Indicators from verbal reports & interviews
(1) Scan claims	"So first I'll just read all my claims on the side"
	"So, I'm just going to read through all of these reasonings and see where I'll begin, how I'll form the map. So I'm just going to read them out loud." (novice)
	SCAN ALL CLAIMS before moving claims: Participant reads aloud every claim presented on initial screen.
(2) Identify & place conclusion	"I think generally, what I always try to do is find whatever I take to be the ultimate conclusion first."
	"All right, okay I'm just thinking - okay yeah that's - that'll be main one, main point."
	PLACE CONCLUSION FIRST: (after reading all the claims) "Okay, so I'm looking for a conclusion, that was, the general use of multimedia increases learning, for students (placed the conclusion first at top)."
	DIRECTIONAL FLOW: Placing the conclusion at top, bottom, right, or left (not center of screen) so that chains of premises flow top-down, left-to-right, or right-to-left, respectively.
(3) Position claims	SORT-FIRST PROCESS: "As I was reading all the reasons, I noticed various similarities. So we're excluding text, we're excluding gratuitous visuals, we're excluding gratuitous sounds. I knew that all of those would typically go up towards the top and then go to support something about decreasing load on some sort of memory, so that helped me at least get that structure set…. I think that I tried to not put any links in until I felt fairly confident that I wanted them there."
	BACKWARD PROCESS: "All right, okay I'm just thinking--that'll be main one, main point…I'm going to begin making a map. I'm going to pick number 11, use of multimedia increases learning. I'm going to bring this over to the section and I'm going to...Okay, let's see. Helps encode into long term memory. I think that definitely relates to-- That's a good reason to have to use multimedia in e-learning…I'm going to add the arrow to connect to use of multimedia increases learning."
	FORWARD PROCESS: "Decrease load on visual working memory, .. Exclude gratuitous visual… Ah.. Personalized communication. I'm thinking … maybe there are corresponding ones to these. I'm thinking

	maybe there are corresponding ones to these. So then, Decrease load on auditory working memory. Yeah, Okay. Then, Decrease load on working visual memory. That's up here".
	DEPTH-FIRST PROCESS: Back-to-back use of backward process or forward process that link three claims into a single chain A → B → C.
	BREADTH-FIRST PROCESS: "I'm trying to figure out... what other examples I can attach to main point, use of multimedia increases learning… I'm going to take this one, exclude gratuitous text…So I'm going to add the arrow to this one and connect [to main conclusion]".
(4) Link claims	"I'm going to add the arrow to connect to use of multimedia increases learning." "So, we'll hook that up there."
	"I feel like that's...obviously that's a negative one. Yeah. It's like, I'm going to add a red arrow to that because negative."
	LINK CLAIM & CHECK CHAIN: Review the whole chain when adding a claim to the chain. "Removing word for word narration, exclude gratuitous sounds will decrease load on auditory working memory… umm.. All those things are gonna help reduce overall cognitive load that students have to deal with…ah…it should help encode into long-term memory."
(5) Review & revise links	RECOGNIZE ERROR: "Use of multimedia increases learning. Increase selective attention. Add async audio narration and animated demonstration with reading text. Okay actually I think this one, increase selective attention. How do I-- oh that's right just delete the arrow."
	BACKWARD PROCESS to review map, "The techniques-the things to do to help encode into long-term memory include personalized communication with the use of pedagogical agents and one technique for doing that would be to add audio narration and animated demonstration."
	FORWARD PROCESS to review map, "Add sync audio narration and animated demonstration with reading text, excluding gratuitous visuals and excluding gratuitous texts will decrease the load on the working visual memory, reduce the cognitive load and encode into long term memory."
	DEPTH-FIRST CHAIN: Review chain of reasoning with backward process
	BREADTH-FIRST PROCESS: "You use multimedia to help encode into long-term memory and reduce overall cognitive load."

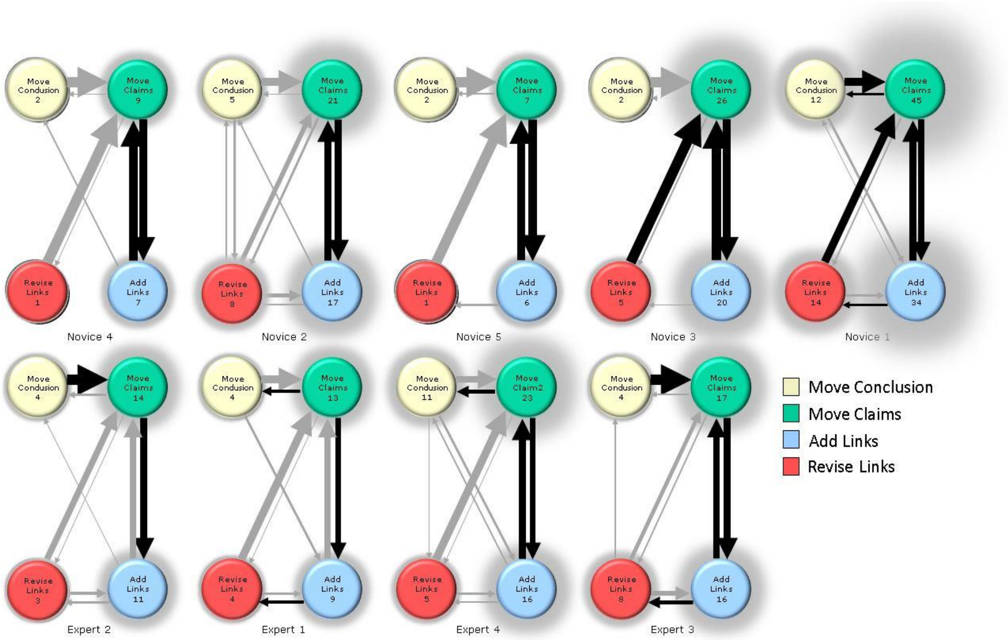
Table 3 shows that some steps (scan claims, revise links) were entirely or almost entirely skipped by one or more participants, and that some occurred later (not earlier) in the process (place conclusion). Steps 3, 4 and 5 were performed iteratively (position claims, link claims, revise links), consistent with what Cañas, Reiska, and Möllits (2017) found and described as the process of "reflective thinking". Which of the iterative processes distinguish high from low performers was determined by sequentially analyzing mapping actions (actions recorded in the jMAP logs) to produce the transitional state diagrams in Fig. 3. The state diagrams, ordered from participants with the lowest to highest map scores, present the relative frequency of action sequences performed by each participant and the relative frequencies that occurred at higher-than-expected frequencies based on z-score tests at $p < .01$.

Table 3
Main steps and strategies with performance measures from logged data and video analysis

	Novices					Experts			
Main Steps & Strategies	1	2	3	4	5	1	2	3	4
Argument map score	14	7	11	5.5	9	19	16	31	26
Step 1) Scan all claims	√		√		√	√	√	√	√
Scan minutes (M = 1.35)	.88	.3	2.8	.65	1.1	1.7	2.5	1.7	.5
Step 2) Place conclusion	na [1]	√	√	√	√	√	√	√	√
#Times conclusion moved	2	5	2	2	12	4	4	11	4
Map format	Bottom	Center	Top	Top	Top	Top	Right	Right	Bottom
Step 3) Position claims	√	√ [2]	√	√	√	√	√	√	√
#Times claim is moved (M = 6.8)	12.4	5.3	6.3	2.9	6.8	4.7	5.8	8.1	9.1
#Claims positioned prior step 2	1	4	0	12	0	1	3	1	3
Backward processing	5	0	27	7	0	17	27	40	0
Forward processing	26	0	0	5	15	2	0	0	25
	31	0	27	12	15	19	27	40	25
Depth-first	3	0	0	0	0	2	0	0	8
Breath-first	17	0	13	3	13	13	25	31	16
	20	0	13	3	13	15	25	31	24
#Moved claims prior to step 4	4	4	5	12	61	18	62	24	13
Step 4) Link claims	√	√	√	√	√	√	√	√	√
Frequency (M = 28.4)	45	35	20	18	31	18	13	49	27
Total links (15 in criterion map)	15	15	14	17	19	14	13	18	14
Change link attributes	3	7	1	0	6	0	6	2	2
Step 5) Revise links	√	√	√	√	√	√	√	√	√
#Links inserted prior	4	9	5	4	1	5	1	12	7
Frequency	31	24	11	11	20	5	2	35	5
#Links added after revision	9	2	2	10	8	1	12	2	5

Note. [1] Never identified the conclusion and placed conclusion at bottom of map as a premise;
[2] None of the strategies apply because all claims were linked directly to the conclusion.

The state diagrams show that 7 of the 9 participants exhibited the tendency to perform steps 3 and 4 (move nodes → add links) in an iterative process. The two participants that did not exhibit this tendency were experts 1 and 2 (with map scores higher than all the novices). Both used the sort-first strategy by moving a large number of claims to their desired location before linking the claims. Secondly, novice 1 and 3 iteratively revised links and moved nodes (suggesting that they did not use the sort-first strategy) when no such tendency was exhibited among the four experts (all with higher map scores). The state diagrams also show how the experts overall exhibit a higher tendency to use a more structured process than the novices - working between moving-conclusion and moving-nodes, moving-nodes and adding-links, and then following adding-links with revising-links. These findings demonstrate that the mapping actions mined by the mapping tool can be used to determine the extent to which students construct their maps using a specific process.



Note. Thickness of arrow conveys relative strength of transitional probability; black arrows identify probabilities that are significantly greater than expected by chance alone based on z-score tests at p < .01 (Bakeman & Gottman, 1997); size of glow emanating from node conveys relative number of times the action was performed.

Fig. 3. Sequential patterns in participants' mapping steps ordered from low to high scorers

Step 1 Scan claims

Analysis of the think-aloud protocols of participants performing step 1 revealed 7 of the 9 participants using the scan-first strategy - scanning and reading all claims (averaging 6.58 seconds per claim) before placing a claim in the map. The remaining participants (two novices with the lowest map scores) started positioning claims before scanning all the claims at 18 to 39 seconds into the mapping task. The mapping tool by itself was unable to determine which participants scanned all claims because no events occur on computer screen to indicate the act of reading a claim. Possible behavioral indicators could be generated by presenting and magnifying the text in each claim when mousing over a claim, tracking mouse cursor movements when the cursor is used to point to each claim as it is being read, and measuring the time elapsed between the time claims are first presented and the time at which a claim is moved.

Step 2 Identify and place conclusion

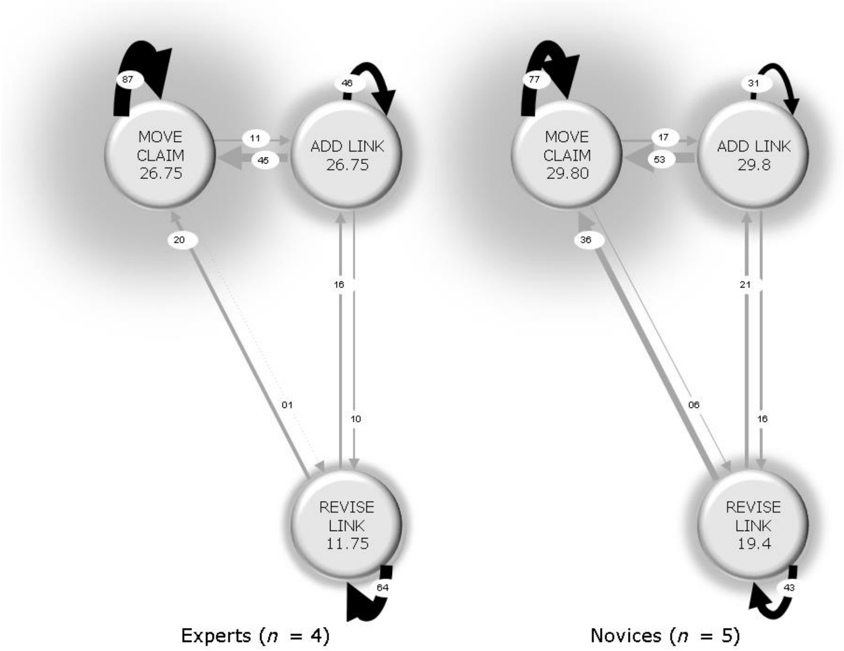
After scanning claims, two experts and three novices used the place-conclusion-first strategy by proceeding (after moving just one or less claims) to place the conclusion into the map. The remaining participants (which included the lowest scoring expert and two lowest scoring novices) moved 3 to 12 claims before placing the conclusion. Furthermore, state diagrams in show (Fig. 3) that all four experts exhibited the tendency to place the conclusion and then move claims (move conclusion → move claims) or vice versa at higher-than-expected frequencies. This suggests that all four experts tried to coordinate the placement of the claims in relation to the conclusion. In contrast, only novice 1 (the novice with the highest map scores) exhibited this behavioral pattern. These findings suggest that identifying and placing the conclusion prior to moving claims and/or coordinating movements between conclusion and claims produces better maps. It produces better maps because argument analysis is a goal-directed task that requires sustained focus on goals while searching for ways to link claims to conclusions (Sharma, Tiwari, & Kelkar, 2012).

The observations also reveal that where the conclusion is placed determines to what extent claims are placed with directional flow or "reading flow" (Aguiar & Correia, 2017). Causal maps with more flow (e.g., more links pointing up than down in a top-down map) have been found to correctly identify more root causes (nodes located at the tail ends of each stem) than maps with less flow (Jeong & Lee, 2012). In this study, the novice with the second lowest score placed the conclusion in mid-screen to produce a hub-and-spoke diagram with no flow - when claims radiate out in all directions from the conclusion resulting in nearly equal numbers of arrows pointing left vs. right and down vs. up. Although hub-and-spoke maps are said to be easier to create, they are more disorienting and require more mental effort to process (Amadieu, Van Gog, Paas, Tricot, & Mariné, 2009). In contrast, two experts in this study (one with the highest map score) placed the conclusion to the far right to create left-to-right flow, making linked claims easier to read and review (Derbentseva & Kwantes, 2014). The mapping tool records each claim's x-y coordinates and can count the number of links pointing up, down, left, and right. But at this time, it cannot identify map orientation to use the counts to measure flow. To do this, the tool will need to track where claims are placed in relation to the conclusion with flow in hub-and-spoke maps, for example, measured in terms of the number of links pointing toward vs. away from the conclusion.

Step 3 Move claims

Sort-first strategy. The process of moving claims was associated with four strategies - sort-first, backward, forward, and breadth-first processing. The experts used the *sort-first strategy* more than the novices to sort claims based on similarities in degree of similarity and generality, as prescribed by Cañas, Reiska, and Möllits (2017, p. 354), while linking few if any claims prior to placing all the claims into their positions. Three novices inserted links after moving 4, 4, 5, and 12 claims - less than the number of claims moved by any of the experts. The state diagrams (Fig. 4) of mapping actions aggregated across all experts vs. all novices confirms this finding. They show that the experts altogether were more likely than the novices to: a) move claims one after another (.87 vs. .77, respectively) while at the same time, were less likely to follow moving a claim by adding a link; b) add one link after another (.46 vs. 31, respectively) once experts began the process of adding links; and c) revise one link after another (.64 vs .43, respectively) with experts making *fewer* number of revisions to links (M = 11.75) than novices (M = 19.4). After revising a link(s), the experts were *less* likely than the novices to go back to moving claims (.20 vs. .36, respectively.

The sort-first strategy produces better maps because it is analogous to sorting jigsaw puzzle pieces prior to piecing them together, simplifying the problem space and making pieces easier to find once needed (Antle, 2013). Placing claims in close proximity increases the number of attempts to find possible links between claims just as eye movement data show that students make significantly more saccades between and more attempts to integrate text and diagrams placed in proximity (Johnson & Mayer, 2012). Furthermore, structural changes can be made without the cumbersome process of undoing links and relinking claims as they are slid up and down a chain or moved from branch to branch. This can help prevent users from entrenching themselves behind their initial maps and/or producing an incomplete map (novice 4 left three isolated clusters of claims). Overall, these findings demonstrate that the mined mapping actions can be used to gauge use of the sort-first strategy.



Note. Values displayed in each step = average number of times the step was performed per participant.

Fig. 4. Sequential patterns in mapping actions between all experts versus all novices

Backward-over-forward processing strategy. Video analysis revealed that the experts used proportionately more backward than forward processing than the novices. Even with the differences in number of experts (n = 4) and novices (n = 5), the backwards process was used to position claims 84 and 39 times by the experts and novices, respectively, whereas the forward process was used 27 vs. 46 times by the experts and novices. This finding is consistent with studies that show how more use of backward than forward processing produces better causal maps (Lee, 2012; Shin & Jeong, 2021). One reason as to why the novices in this study showed the tendency to used proportionately more forward processing is that people tend to use forward processing by default to evaluate arguments when faced with unfamiliar problems (Heit, 2007; Oaksford & Hahn, 2007) – in this case, people with no prior training in analyzing arguments. The relationship between group and approach was statistically significant, $\chi2(1, N = 196) = 18.28$, $p < .001$.

Modifications to the mapping tool can be made to enable it to assess where a newly moved node is placed in relation to the most recently moved node to determine whether a student is using backward processing (places node below node in top-down map) or forward processing (places node above node in top-down map). To take map format into account (e.g., spoke-and-hub, top-down map), the mapping tool must also determine the relative position of the most recently moved node in relation to the location of the conclusion. In contrast, backward processing in a left-to-right map is indicated by the act of moving a node placed to the left of the most recently moved node which itself is placed to the immediate left of the conclusion.

Breadth-over-depth-first strategy. The experts and novices exhibited a tendency to use breadth-first processing to link claims. The video analysis revealed that each expert and novice used breadth-first processing to position claims an average of 21.25 and 9.2 times, respectively. In contrast, each expert and novice used depth-first processing an average of 10 and 3 times. The relationship between group and process used was not statistically significant, $\chi 2(1, N = 144) = .321$, $p < .571$. Considering that the criterion map can be completed using breadth-first versus depth-first processing 8 versus 5 times at minimum, the numbers still show both experts and novices exhibiting the tendency to use more breadth-first than depth-first processing. This finding is consistent with the claim that breadth-first is the process of choice when working top-down from given goals (Sharma, Tiwari, & Kelkar, 2012). Also, the finding that experts used breadth-first processing more than the novices is also consistent with Shin and Jeong's (2021) study which found that students producing better causal maps used proportionately more breadth- than depth-first processing.

At this time, no studies provide findings to explain why breadth-first processing produces better maps. One possible explanation is that using the sort-first strategy (which may be associated with higher map scores) involves placing claims sharing the same level of generality side by side (in a top-down map). To test the efficacy of using breadth-first processing with a large sample of participants, the mapping tool requires further development so that it can: a) perform the similar functions used to identify backward and forward processing (as explained above); b) test and specify the optimal range of angles used to determine when a node is placed to the right, left, above, and below another claim given the length-width dimensions of the claim; c) adjust measures in relation to the minimum number of breadth-first and depth-first processes required to complete the criterion map; and d) ignore claims movements used merely to make cosmetic changes - which occur more often in the latter half of the map construction process.

Step 4 Link claims

The experts and novices inserted links to connect claims an average of 26.75 and 29.8 times, respectively (nearly double the 15 total links in the criterion map). The expert and novices' final maps contained an average of 14.75 and 16 total links, respectively. These numbers show that both the experts and novices ultimately inserted double the number of links than necessary, requiring them at some point in time to delete links when errors were identified. Each expert and each novice changed the attributes of a link (from positive to negative) an average of 2.50 versus 3.40 times, respectively, while the criterion map contained only one link that required its valence to be changed from positive to negative. Overall, the findings do not reveal notable patterns that distinguish the experts from the novices because most of the work of identifying the relationships between claims is reflected in the act of positioning a claim to the left, right, above, or below another claim (using the strategies such as backward, forward, depth-first, and bread-first processing), not in the act of inserting a link between two claims.

However, video analysis revealed that the moment after inserting a link to connect two claims, participants traced through and reviewed the series of claims lying upstream and/or downstream from the newly linked claim to assess its flow in logic. No observable actions performed on the computer screen could be mined and used by the mapping tool to quantify when and how often the participants perform this action. This action could potentially be detected by the mapping tool if the tool were to: a) magnify the text presented in each claim to induce users to mouse over each claim as they read and trace through linked claims (as mentioned previously); and b) record the movement of the mouse cursor as it is being used to coordinate and direct attention to each claim while tracing through the map.

Overall, no notable differences were found between experts and novices in how links were added to maps. This suggests that mining and analyzing the act of linking claims by itself provides no process data of strategic value. However, identifying the first act of linking two claims immediately after the initial movement of claims is necessary to assess use of the sort-first strategy (as discussed above). Furthermore, strategic data can be obtained by enabling the mapping tool to: a) prompt users to identify a mediating claim X that explains why claim A supports claim B (A → X → B) each time users insert a link; and b) analyze inserted links to efficiently determine to what extent users are creating directional flow (as described above). On the flipside, the mapping tool could be modified to let users skip this step entirely by letting the mapping tool auto-insert links the moment a claim is placed adjacent to another claim, but doing so in a way that does not make re-positioning claims to explore tentative relationships and to make structural changes overly cumbersome. The mapping tool can do this by unlinking a claim B by simply dragging B away from its neighbouring claims and automatically relinking orphaned claims to change A→B→C to A→C.

Step 5 Review and revise links

Analysis of the mined mapping actions revealed that the expert and novice with the highest score within group performed the highest number of revisions to map links (35 and 31, respectively). The expert and novice with the lowest score within group performed the least number of revisions (2 and 11, respectively). Hence, the correlation between scores and number of revised links was .82 among the experts and .53 among the novices. Overall, the experts performed fewer revisions to links on average (M = 11.75) than the novices (M = 19.4). These patterns suggest that map scores improve in proportion to effort invested in finding and correcting errors in linked claims – a finding that is consistent with prior studies that find students construct better maps when they perform more frequent and more complex revisions to their maps (Shin & Jeong, 2021). The observed patterns also suggest that fewer corrections are needed and/or performed when users possess prior knowledge on how to analyze arguments.

A combination of performance data on mapping actions can be used to identify users struggling with their maps. Table 4 shows that expert 2 (lowest scorer among the experts) and novice 4 (lowest scorer among the novices) moved the highest and second highest number of claims within their group (62 and 12, respectively) before inserting the first link. Also, expert 2 inserted only 1 link (fewest among the experts) prior to making the first revision to a link. Similarly, novice 4 inserted 4 links (tied for second fewest among the novices) prior to making the first revision to a link. Expert 2 then inserted 12 more links (most among the experts) after making the last revision to a link while novice 4 inserted 10 more links (most among the novices) after making the final revision to a link. The mapping tool can compute the norms on these three metrics and cross-index these measures with the user's current map score based on present links and links implied by position and proximity between nodes.

Finally, video analysis revealed that backward, forward, depth-first, breadth-first processing was used to visually trace through linked claims to search for errors. As discussed above, no observable actions were performed on screen to enable the mapping tool to mine this behavior and to quantify when and how often participants use each strategy to review links. Again, these actions could potentially be detected by inducing users to mouse over claims to magnify and read the claims, and record cursor movements as users point to and trace through linked claims with the cursor.

3.3. Cognitive acts denoted by discrete mapping actions

The thematic analysis of the think-aloud protocols revealed 14 verbalized-cognitive acts (Table 4) associated with 10 discrete mapping actions performed on the jMAP screen. One mapping action (move claim next to previously moved claim) served as a behavioral indicator for three closely similar verbal-cognitive actions: identifying level of claim (verbalizations with words like "over", "below", "under", "up", identifying association between claims ("related to", "involves"), and identifying causal relationship between claims ("results in", "helps"). The four cognitive acts that were found to have no corresponding mapping action to denote their presence were read/scan claims, interpret/comprehend a claim, review chain of reasoning, and recognize an error. Furthermore, the cognitive act of justifying or explaining the reason for linking two claims was not observed in any of the think-aloud protocols. The participants simply re-stated that claim B is true "because of" claim A without identifying the mechanism or the intermediate claim that link the two claims (e.g., reduce cognitive load → improve information processing in working memory → increase encoding into long term memory). Performing this action requires deeper knowledge of learning theory, which was a domain outside of the participants' area of expertise except for one expert. Furthermore, the participants were not prompted to explain the mechanism, were not given the option to insert additional claims into their maps and were not instructed to annotate the links with explanations.

Of the 10 cognitive actions that were denoted by discrete mapping actions, no data is available to determine precisely how often students performed each of the 10 actions covertly without exhibiting it overtly via a mapping action (e.g., finding a claim as opposed to placing the cursor on a claim, and contemplating the relationship of one claim to another claim as opposed to moving and placing the claim next to the related claim). However, the verbal protocol transcript revealed that the researcher prompted the novices to talk aloud (when the participants stopped talking after 5 seconds passed without talk) a total of 5 times (3 times for novice 1, and once for novices 2 and 3) and 3 times with the experts (twice for expert 1, and once for expert 3) relative to the total of 510 and 680 cognitive acts verbalized by the novices and experts, respectively. The participants' frequent use of the cursor to point to claims as they read each claim out loud (82%, 83%, 43%, 39%, and 44% of all coded cognitive acts performed the novices, and 33%, 48%, 3%, and 25% of all the coded cognitive acts performed by the experts) provided a good indication that most of the cognitive acts were verbalized by the participants as the constructed their maps.

Overall, these findings show that the actions mined by the mapping tool can capture 10 of the 15 cognitive acts (or 66% a most) used to construct a map. As a result, the mapping action alone (without verbal reports) can be analyzed to identify the cognitive acts used to produce better maps. Detecting the cognitive acts that have no associated mapping actions could be achieved by adding new functions to the mapping tool. For example, some ways to detect the process of reading claims is to: 1) gauge the lag time separating the previous action and following action; and 2) initially position all claims in a holding area, display these claims only when participants move the mouse cursor into the holding area, and hide the claims once the mouse cursor is moved out of the holding area. To detect the process of tracing and reviewing chained claims, magnifying the text in each claim as students' mouse over each claim.

Table 4
Cognitive actions in think-aloud protocols with actions observed and mined in mapping tool

Verbalized cognitive act	Mapping action observed on screen
Read claim	*None*
Identify main conclusion	Position conclusion in the map
Interpret/comprehend a claim	*None*
Identify level of claim	Move claim next to previously moved claim
Identify association between claims	Move claim next to previously moved claim
Identify causal relationship between claims	Move claim next to previously moved claim
Identify the dependency of two claims	Connect arrows to node at same point
Identify independency of two claims	Connect arrows to node at 2 different points
Identify irrelevant claim	Claim is not connected or placed to the side
Identify negative association	Change attribute of a link
Make a cause-effect relationship	Insert link to connect two claims
Review chain of reasoning	*None*
Recognize error in reasoning	*None*
Correct error in reasoning	Delete link, Re-position node & Re-link

3.4. Implications for map tool design and instruction

The purpose of this study was to identify strategies experts and novices (with and without prior knowledge on how to analyze arguments) used to construct better argument maps and to identify how mapping actions can be mined and analyzed in mapping tools to measure student use of the mapping processes and strategies. Through comparative analysis of the findings from the videos, think-aloud protocols, retrospective interviews, and mined mapping actions, this study validated the 5-step process for constructing maps as prescribed by Aguiar and Correia (2017). This study also identified several strategies used to perform each step and provides further validation of findings from prior studies that show backward processing and breadth-first processing to be associated with and possibly contribute to higher maps scores (Lee, 2012; Shin & Jeong, 2021). Most of all, this study provides numerous recommendations (as described above) on how mapping tools can be refined to mine and analyze on-screen behaviors to measure how students are using the specific strategies to complete their maps or to complete any kind of learning or problem-solving task that involve the process of visually and spatially organizing conceptual components into larger structures (e.g., visual design, root cause analysis).

Such refinements enable the mapping tools to be used to conduct large-scale studies to determine how and which strategies improve students' map scores and reduce variance in students' scores. Mapping tools can then be further modified to deliver interventions to support the use of proven strategies, test for alignment between intended processes and what students actually do, and identify new behaviors and strategies.

Although the 5-step process and the specific strategies identified in this study must still be tested for predictive validity, these tentative findings provide a preliminary glimpse into some of the possible strategies class-room instructors can use to help students analyze arguments more effectively. For example, instructors can help by: 1) explicitly instructing students to begin the task by finding the conclusion and placing it at the top of the screen (not center screen); 2) providing students with a mapping tool that enable them to freely move and place claims on screen without having to link a claim the moment it is placed on screen; 3) instructing students to find claims that are similar in degree of generality and specificity and line up similar claims into rows below the conclusion; and 4) progressively position and link claims working backwards from the conclusion with all links pointing towards the conclusion to facilitate the process of reviewing and identifying errors in the way the claims are linked and chained. To help identify students who are struggling the most as they are constructing their maps, instructors can look out for students who deviate from the processes prescribed above. Furthermore, these very same processes and strategies can be taught and used to help students construct better causal maps to decompose problems and identify root causes and better concept maps to organize or break down semantic and hierarchical relationships between concepts.

3.5. Limitations and directions for future research

Future studies can address some of the limitations noted below and test specific inferences raised in this study by doing the following: 1) develop, test, and refine algorithms in the mapping tool using larger data sets to produce metrics that best predict map scores; 2) measure each mapping action with more specificity (e.g., add correct vs. incorrect link) and strategy (e.g., place correctly identified conclusion at top of map) in reference to a criterion map that vary in structure (e.g. more depth than breadth) and adjust measures to take differences in map structure into consideration; 3) develop and test more precise ways to reveal internal cognitive processes (e.g. scanning claims) on screen to expand the range of behaviors that can be mined and analyzed; 4) measure, verify, and determine to what extent the use of each strategy and required mental effort is dependent on prior knowledge of argumentation analysis and prior content knowledge; 5) use eye tracking systems to examine how mapping with versus without auto-linking affects the extent to which students use, for example, the sort-first strategy (or more specifically, the number of eye saccades between grouped claims), the rate at which grouped claims are correctly vs. incorrectly linked and the resulting number of times students revise links; and 6) determine the minimum number of claim moves performed at the start of the mapping process to produce measures of breadth-first vs. depth-first processing that best predict map scores – measures that are not confounded by minor actions used to make cosmetic changes to the maps.

4. Conclusions

This study found that experts applied a more sequential application of a five-step process (scan claims, identify and place conclusion, move and position claims, link related claims, review and correct links between claims) to produce more accurate argument maps than novices. The novices showed the tendency to position and link claims as a joint action, making map revision more cumbersome. The experts exhibited the tendency to work backward from conclusion to claim while the novices exhibited the tendency to work forward from claims to conclusion. The findings identify processes that differentiate experts from novices and identify strategies that can be potentially taught and used by students to improve map quality.

Overall, this study reveals complexities and nuances in the processes of constructing argument maps and some of the factors to be taken into consideration when choosing the most effective processes. Developing mapping tools that can mine mapping actions and produce analytics like those demonstrated in this study may help advance and build on prior research aimed at identifying, measuring, modeling, and empirically testing strategies used to construct not only argument maps, but also concept maps and causal diagrams. Such research can potentially increase our understanding of what and how specific strategies impact student learning, understanding, and performance in problem-solving. Operationalizing the observed strategies and embedding them into next-generation mapping tools can potentially help students successfully analyze complex problem domains with better control over differences in critical thinking and reasoning abilities.

Author Statement

The authors declare that there is no conflict of interest.

ORCID

Allan Jeong https://orcid.org/0000-0001-5797-3223

Hae Young Kim https://orcid.org/0000-0002-5001-8431

References

Aguiar, J. G., & Correia, P. R. M. (2017). From representing to modelling knowledge: Proposing a two-step training for excellence in concept mapping. *Knowledge Management & E-Learning, 9*(3), 366–379.

Al-Ajlan, A. (2015). The comparison between forward and backward chaining. *International Journal of Machine Learning and Computing, 5*(2), 106–113.

Amadieu, F., Van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction, 19*(5), 376–386.

Antle, A. N. (2013). Exploring how children use their hands to think: An embodied interactional analysis. *Behaviour & Information Technology, 32*(9), 938–954.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. Cambridge University Press.

Beitz, J. M. (1998). Concept mapping: Navigating the learning process. *Nurse Educator, 23*(5), 35–41.

Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From design to implementation to practice a learning by teaching system: Betty’s Brain. *International Journal of Artificial Intelligence in Education, 26*(1), 350–364.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Cañas, A. J., Novak, J. D., & Reiska, P. (2015). How good is my concept map? Am I a good Cmapper? *Knowledge Management & E-Learning, 7*(1), 6–19.

Cañas, A. J., Reiska, P., & Möllits, A. (2017). Developing higher-order thinking skills with concept mapping: A case of pedagogic frailty. *Knowledge Management & E-Learning, 9*(3), 348–365.

Chen, J. C., Li, Y., & Shady, B. D. (2010). From value stream mapping toward a lean/sigma continuous improvement process: An industrial case study. *International Journal of Production Research, 48*(4), 1069–1086.

Chen, J., Wang, M., Dede, C., & Grotzer, T. A. (2021). Analyzing student thinking reflected in self-constructed cognitive maps and its influence on inquiry task performance. *Instructional Science, 49*(3), 287–312.

Clark, R. (2002). Six principles of effective e-learning: What works and why. *The e-Learning Developer's Journal*, 6(2), 1–10.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational Psychological Measurement, 20*(1), 37–46.

Correia, P. R. M., & Aguiar, J. G. (2014). Concept mapping informed by cognitive load theory: implications for tasks involving learner-generated Cmaps. In *Proceedings of the sixth International Conference on Concept Mapping* (pp. 150–157). Santos, Brazil.

Cottrell, S. (2017). *Critical thinking skills: Effective analysis, argument and reflection*. Bloomsbury Publishing.

Davies, M. (2011). Introduction to the special issue on critical thinking in higher education. *Higher Education Research & Development, 30*(3), 255–260.

Davies, M., Barnett, A., & van Gelder, T. (2019). Using computer-assisted argument mapping to teach reasoning to students. In J. A. Blair (Ed.), *Studies in Critical Thinking* (Vol. 8, pp. 131–176). Windsor, Canada: Windsor Studies in Argumentation.

Derbentseva, N., & Kwantes, P. (2014). Cmap readability: Propositional parsimony, map layout and semantic clarity and flow. In *Proceedings of the Sixth International. Conference on Concept Mapping* (Vol. 2, pp. 86–93). Santos, Brazil.

Easterday, M. W., Aleven, V., & Scheines, R. (2007). Tis better to construct than to receive? The effects of diagram tools on causal reasoning. In *Proceedings of the 13th International Conference on AI in Education* (pp. 93–100). IOS Press.

Easterday, M. W., Aleven, V., Scheines, R., & Carver, S. M. (2009). Constructing causal diagrams to learn deliberation. *International Journal of Artificial Intelligence in Education, 19*(4), 425–445.

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.

Ennis, R. H. (1982). *Identifying implicit assumptions. Synthese, 51*(1), 61–86.

Eppler, M. J. (2006). A comparison between concept maps, mind maps, conceptual diagrams, and visual metaphors as complementary tools for knowledge construction and sharing. *Information Visualization, 5*(3), 202–210.

Green, N. L., Branon, M., & Roosje, L. (2019). Argument schemes and visualization software for critical thinking about international politics. *Argument & Computation, 10*(1), 41–53.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp.105–117). Thousand Oaks, CA: Sage

Harrell, M. (2007). Using argument diagramming software to teach critical thinking skills.

In *Proceedings of the 5th International Conference on Education and Information Systems, Technologies and Applications*.

Harrell, M. (2011). Argument diagramming and critical thinking in introductory philosophy. *Higher Education Research & Development, 30*(3), 371–385.

Heit, E. (2007). What is induction and why we study it? In A. Feeney & E. Heit (Eds.), *Inductive Reasoning: Experimental, Developmental, and Computational Approaches* (pp.1–24). Cambridge University Press.

Hinkelmann, K. (2004). *Forward chaining vs. backward chaining*. University of Applied Sciences Northwestern Switzerland, School of Business. Retrieved from http://knut.hinkelmann.ch/lectures/KE2011/KE-4_FC_vs_BC.pdf

Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2010). Effects of representational guidance during computer-supported collaborative learning. *Instructional Science*, *38*(1), 59–88.

Jeong, A. C. (2020). Developing computer-aided diagramming tools to mine, model and support students' reasoning processes. *Educational Technology Research and Development, 68*(6), 3353–3369.

Jeong, A. C., & Lee, W. J. (2012). Developing causal understanding with causal maps: The impact of total links, temporal flow, and lateral position of outcome nodes. *Educational Technology Research & Development, 60*, 325–340.

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, *18*(2), 178–191.

Kinchin, I. M. (2001). If concept mapping is so helpful to learning biology, why aren't we all doing it? *International Journal of Science Education, 23*(12), 1257–1269.

Körner, C. (2005). Concepts and misconceptions in comprehension of hierarchical graphs. *Learning and Instruction, 15*(4), 281–296.

Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*(5), 1245–1260.

Larson, A. A., Britt, M. A., & Kurby, C. A. (2009). Improving students' evaluation of informal arguments. *The Journal of Experimental Education*, *77*(4), 339–366.

Larson, M., Britt, M. A., & Larson, A. A. (2004). Disfluencies in comprehending argumentative texts. *Reading Psychology*, *25*(3), 205–224.

Lee, W. J. (2012). *Exploring students' mapping behaviors and interactive discourses in a case diagnosis problem: Sequential analysis of collaborative causal map drawing processes*. Doctoral dissertation, Florida State University, USA.

Nielsen, J. (1994). Estimating the number of subjects needed for a thinking aloud test. *International Journal of Human–Computer Studies, 41*(3), 385–397.

Oaksford, M., & Hahn, U. (2007). Induction, deduction, and argument strength in human reasoning and argumentation. In A. Feeney & E. Heit (Eds.), *Inductive Reasoning: Experimental*, *Developmental, and Computational Approaches* (pp. 260–301). Cambridge University Press.

Roksa, J., & Arum, R. (2011). The state of undergraduate learning. *Change: The Magazine of Higher Learning, 43*(2), 35–38.

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2017). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review, 30*(2), 431–455.

Serrat, O. (2017). The five whys technique. In O. Serrat (Ed.), *Knowledge Solutions: Tools, Methods, and Approaches to Drive Organizational Performance* (pp. 307–310). Springer.

Sharma, T., Tiwari, N., & Kelkar, D. (2012). Study of difference between forward and backward reasoning. *International Journal of Emerging Technology and Advanced Engineering*, *2*(10), 271–273.

Shin, H. S., & Jeong, A. (2021). Modeling the relationship between students' prior knowledge, causal reasoning processes, and quality of causal maps. *Computers & Education*, *163*: 104113.

Sweller, J., Clark, R. E., & Kirschner, P. A. (2011). Teaching general problem solving does not lead to mathematical skills or knowledge. *EMS Newsletter* (March), 41–42.

Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences*, *12*(2), 183–218.

Terjesen, S., & Patel, P. C. (2017). In search of process innovations: The role of search depth, search breadth, and the industry environment. *Journal of Management*, *43*(5), 1421–1446.

ThinkReliability. (2007). *Root cause analysis videos*. Retrieved from http://www.thinkreliability.com/Root-Cause-Analysis-Video-library.aspx

Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.

Tversky, A., & Kahneman, D. (1980). Causal schemata in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in Social Psychology* (pp. 49–72). Erlbaum.

Twardy, C. (2004). Argument maps improve critical thinking. *Teaching Philosophy, 27*(2), 95–116.

van Bruggen, J. M., Boshuizen, H. P. A., & Kirschner, P. A. (2003). A cognitive framework for cooperative problem solving with argument visualization. In P. A. Kirschner, S. J. Buckingham Shum, & C. S. Carr (Eds.), *Visualizing Argumentation* (pp. 25–47). Springer.

van Bruggen, J. M., Kirschner, P. A., & Jochems, W. (2002). External representation of argumentation in CSCL and the management of cognitive load. *Learning and Instruction*, *12*(1), 121–138.

van den Braak, S. W., van Oostendorp, H., Prakken, H., & Vreeswijk, G. A. W. (2006). A critical review of argument visualization tools: Do users become better reasoners? In *Proceedings of the Workshop notes of the ECAI-06 Workshop on Computational Models of Natural Argument* (pp. 67–75).

van Drie, J., van Boxtel, C., Jaspers, J., & Kanselaar, G. (2005). Effects of representational guidance on domain specific reasoning in CSCL. *Computers in Human Behavior, 21*(4), 575–602.

van Gelder, T. (2007). The rationale for Rationale™. *Law, Probability and Risk, 6*, 23–42.

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). New York, NY: Palgrave Macmillan.

Wang, S. (2019). *Providing intelligent and adaptive support in concept map-based learning environments*. Doctoral dissertation, Arizona State University, USA.

Weinerth, K., Koenig, V., Brunner, M., & Martin, R. (2014). Concept maps: A useful and usable tool for computer-based knowledge assessment? A literature review with a focus on usability. *Computers & Education, 78*, 201–209.

Wu, B., & Wang, M. (2012). Integrating problem solving and knowledge construction through dual mapping. *Knowledge Management & E-Learning, 4*(3), 248–257.

Wu, B., Wang, M., Kirschner, P. A., & Spector, J. M. (2018). Using cognitive mapping to foster deeper learning with complex problems in a computer-based environment. *Computers in Human Behavior, 87*, 450–458.

Wright, W., Sheffield, D., & Santosa, S. (2017). Argument mapper: Countering cognitive biases in analysis with critical (visual) thinking. In *Proceedings of the 21st International Conference Information Visualisation* (pp. 250–255). IEEE.