# Modeling the Processes of Diagramming Arguments that Support and Inhibit Students' Understanding of Complex Arguments

**Abstract:** Research on the efficacy of diagramming complex arguments has been mixed. One reason for the mixed findings is that the precise processes students use to construct argument diagrams have yet to be fully examined. This study identified sequential patterns in argument diagramming actions performed by graduate-level students that created high versus low quality diagrams. Transitional state diagrams revealed patterns in the action sequences used by high vs. low performing students - processes believed to either help or inhibit students' ability to construct accurate argument diagrams and achieve a better understanding of complex arguments. The findings reveal processes that can be embedded into future diagramming software to test how particular processes affect students' analysis and understanding of complex arguments.

## Purpose

Research on the efficacy of using visual diagramming tools to facilitate argument analysis has been mixed (Braak et al., 2006; Ruiz-Primo & Shavelson, 1996). One reason for the mixed findings is that empirical studies have yet to be conducted to formally identify the sequential steps and reasoning processes students use when constructing argument diagrams. This study developed a set of visual analytic software tools to record, sequentially analyze, visualize, identify and compare sequential patterns in argument diagramming actions performed by graduate-level students that created high and low quality argument diagrams. Transitional state diagrams were then created and used to visualize sequential patterns found in the actions of students that created high and low quality argument diagrams. The transitional state diagrams were then compared to reveal action sequences that were used by high performing students, but not used by low performing students and vice versa. The unique action sequences performed only by the high performing students can reveal the types of diagramming processes that can be promoted and scaffolded to help students construct more accurate argument diagrams and improve understanding of complex arguments.

## Introduction

Critical thinking is an important skill that enables one to accurately reason and judge information and become lifelong learners for the 21st century. It has been defined as 'the art of analyzing and evaluating thinking with a view to improving it' (Paul & Elder, 2001) and an intellectual standard that includes clarity, accuracy, precision, relevance, depth, logic, and breadth (Mclean, 2005). However, recent research suggests that many college students fail to develop critical thinking skills to the extent that they can effectively use them (Kuhn, 1991). To address this problem, various methods have been used to teach students the skills of argumentation and argument analysis across many disciplines. Argument analysis is the study of logical relationships between propositions presented in an argument (which can be mutually supporting or opposing opinions/claims) in order to reason through premises to reach a conclusion. In argument analysis, students identify the functional roles of each proposition (i.e., conclusion, premise, co-premise, counterargument), analyze the hierarchical relationship among propositions (i.e., levels of premise), and evaluate the quality and line of reasoning. This process helps students to correctly judge the quality and identify flaws within an argument and help students to make well-reasoned decisions.

Given that arguments are often complex and ill-structured, argument-mapping software like Belvedere (De Neys, 2006) and Rationale (van Gelder, 2007) have been developed to help students visualize/identify hierarchical relationships between minor/major premises and claims (Braak, 2006). Some diagramming software, like REASON (ThinkReliability, 2007), prescribe the use of specific logic rules and processes such as the backward reasoning or goal-driven approach (Sharma, 2012). Yet, a critical review of the research on argument diagramming/visualizing tools revealed that the majority of the studies found no significant differences (Braak at al., 2006) and/or were flawed in design. Furthermore, students' maps often varied widely in accuracy regardless of the instructional intervention (Scavarda et al., 2006). Ruiz-Primo and Shavelson's (1996) review of the research lead to the conclusion that students' diagrams should not be used to assess learning until students' facility, prior knowledge, and processes used to create the diagrams are thoroughly examined.

Given that no studies at this time have modeled, identified, and/or validated prescribed mapping processes that enable/inhibit students' to accurately analyze complex arguments, Author (2010) created jMAP to chronologically log each action a student performs while constructing an argument diagram in jMAP. This data can be sequentially analyzed to identify the processes used by students to produce high and low quality argument diagrams. In particular, processes associated with informal reasoning fallacies performed while constructing argument diagrams can in theory be detected by observing diagramming processes. For example, a hasty generalization or leap to conclusion can be observed when a student creates a link (D→A) when the effects of D on A is mediated by B (D→B→A). Circular reasoning can be observed when a student links A→B and B→A. To date, no studies have examined the diagramming/actions sequences associated with reasoning fallacies (including actions that immediately precede/follow such action sequences) in the course of constructing argument diagrams. Using this approach, this study addressed the following questions:

1. What sequential patterns in students' diagramming actions produce the most versus least accurate argument diagrams?
2. What are the differences in processes used to produce the most versus least accurate argument diagrams?

## Method

Seventeen graduate students in an online graduate-level course on computer-supported collaborative learning at a large Southeastern university reviewed arguments produced by students in an online debate (but from another course) to support/oppose the claim: "One's choice of media significantly affects learning". After viewing a video on how to use jMAP, students downloaded a jMAP file to diagram the supporting arguments and another jMAP file to diagram opposing arguments. In both cases, students were presented an initial screen (Figure 1) containing nodes that represented the main claim and supporting/opposing premises.

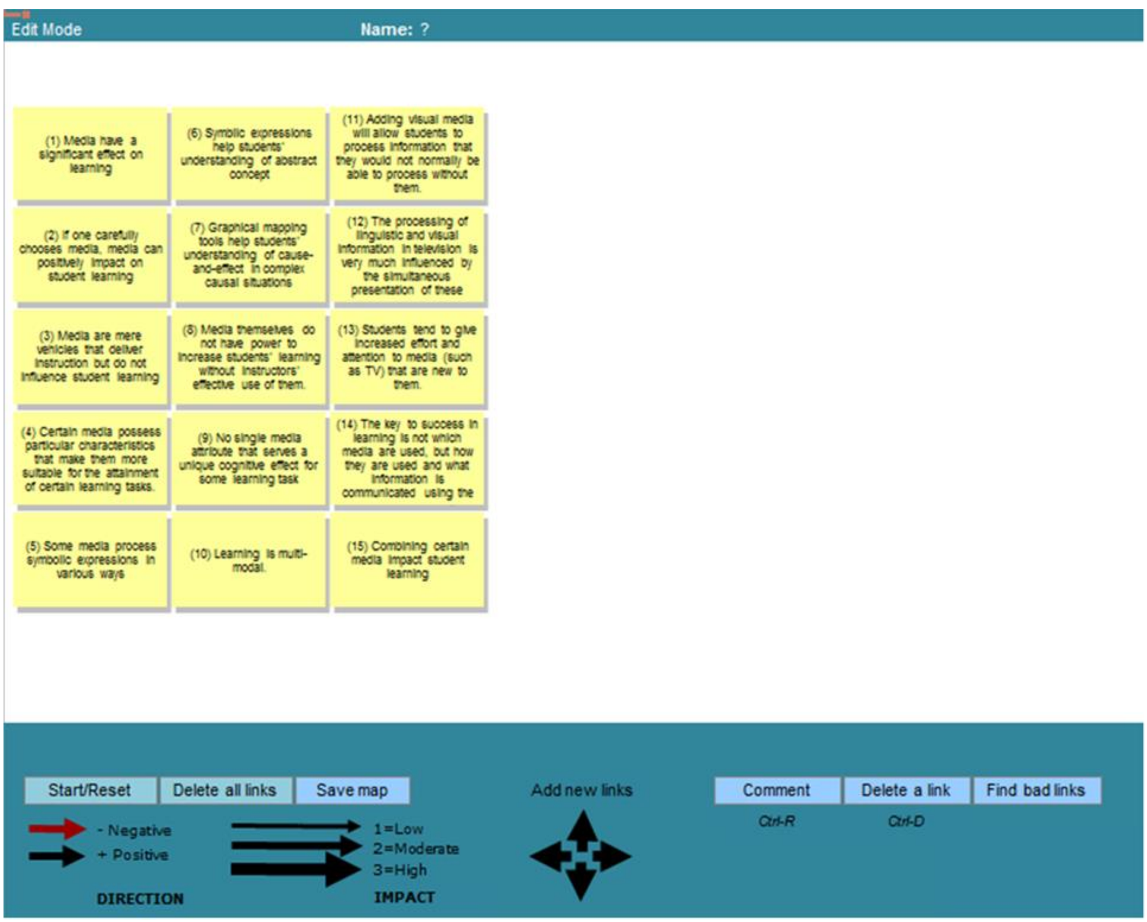
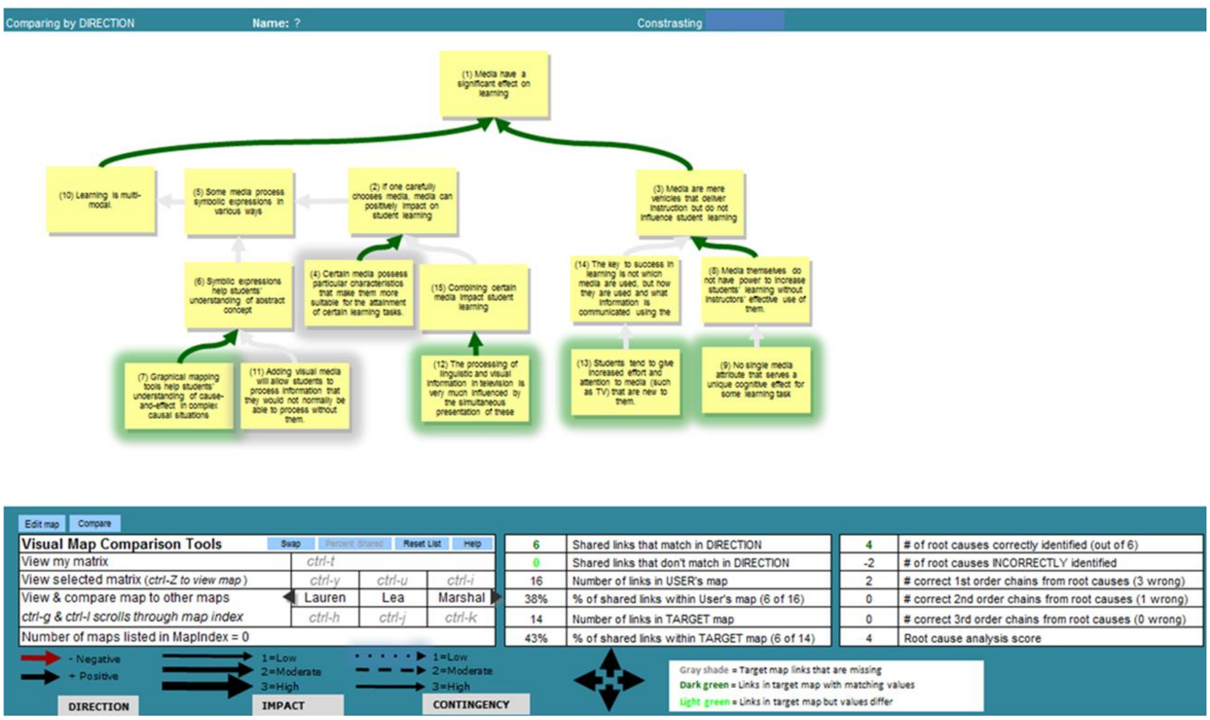



Figure 1. Students' initial screen with premises randomly arranged on the screen.

The students' diagrams were imported into a jMAP file containing the instructors' argument diagram (Figure 2) to score the students' diagrams on three criteria presented in Figure 3: a) percentage of links that match those in the instructor's map; b) number of nodes correctly identified as a root premise (node with no arrows pointing inward); and c) number of correct links stemming directly from each correctly identified root premise up to the main claim. Using a cumulative score to identify diagrams with high versus low accuracy, data from the top 6 and bottom 6 diagrams of the *supporting* arguments and from the top 6 and bottom 6 diagrams of the *opposing* arguments were selected to produce 12 best and 12 worst diagrams. The logged actions (Figure 4) recorded in jMAP from the bottom diagrams were aggregated and then reduced down into six categories to capture more general patterns in students' actions. The same process was repeated with the top diagrams.

The logged actions from the bottom diagrams were imported into the Discussion Analysis Tool (DAT) to produce the frequency, transitional probability, and *z*-score matrices (Figure 5). The frequency matrix shows for example that when these students added a link, 40 of the 129 actions that immediately followed were to add yet another link. To determine if this transitional probability of .31 was significantly higher/lower than expected probabilities (and whether AddLink→AddLink can be deemed to be a sequential "pattern"), the *z*-score matrix shows for this particular action-action sequence a *z*-score of 6.55 (which is greater than the critical *z*-score of ±1.96 at $p < .05$). As a result, the AddLink→AddLink sequence was found to be a sequential pattern in the actions used to produce the bottom diagrams. DAT converted the probabilities into the right transitional state diagram (Figure 6). This process was repeated with data from the top diagrams to produce the left transitional state diagram.




*Note*: Dark/gray colored arrows identify links present/missing in student x's diagram; Nodes with green halos identify lowest level premises correctly identified by student x.

Figure 2. Instructor's diagram visually and quantitatively compared with student x's diagram.

| Close | Shared links that match in DIRECTION | % of shared links within User's map | # of minor premises correctly identified | # of correct 1st order links from minor premises | # of correct 2nd order links from minor premises | # of correct 3rd order links from minor premises | Score |
|---|---|---|---|---|---|---|---|
| [illegible] | 6 | 37.5% | 4 | 2 | 0 | 0 | 97.8 |
| [illegible] | 4 | 22.2% | 4 | 1 | 0 | 0 | 66.2 |
| [illegible] | 5 | 20.8% | 0 | 0 | 0 | 0 | 62.1 |
| [illegible] | 5 | 10.9% | 0 | 0 | 0 | 0 | 61.1 |
| [illegible] | 3 | 18.8% | 5 | 1 | 0 | 0 | 56.9 |
| [illegible] | 2 | 20.0% | 4 | 1 | 1 | 0 | 56.0 |
| [illegible] | 2 | 14.3% | 4 | 1 | 0 | 0 | 45.4 |
| [illegible] | 3 | 16.7% | 0 | 0 | 0 | 0 | 41.7 |
| [illegible] | 2 | 15.4% | 3 | 0 | 0 | 0 | 34.5 |
| [illegible] | 2 | 14.3% | 3 | 0 | 0 | 0 | 34.4 |
| [illegible] | 2 | 14.3% | 1 | 0 | 0 | 0 | 32.4 |
| [illegible] | 1 | 12.5% | 3 | 0 | 0 | 0 | 24.3 |
| [illegible] | 1 | 7.1% | 3 | 0 | 0 | 0 | 23.7 |
| [illegible] | 1 | 7.1% | 2 | 0 | 0 | 0 | 22.7 |
| [illegible] | 1 | 12.5% | 1 | 0 | 0 | 0 | 22.3 |
| [illegible] | 1 | 6.7% | 0 | 0 | 0 | 0 | 20.7 |
| [illegible] | 0 | 0.0% | 0 | 0 | 0 | 0 | 10.0 |

Figure 3. Screen shot from jMAP showing scores assigned to students' diagrams of supporting arguments.

| Category | Code | Definition |
|---|---|---|
| LINK | ADDR | added new link pointing to the right |
| | ADDL | added new link pointing to the left |
| | ADDU | added new link pointing up |
| | ADDD | added new link pointing down |
| | LK2 | attached link to the affected node |
| RELINK | RLK1 | redirected the existing link to a new causal node |
| | RLK2 | redirected the existing link to a new affected node |
| - | ULK1 | detached the beginning tail of the link |
| - | ULK2 | detached the end of the link |
| ATTR | ATT- | changed link to color red to convey a negative or inverse relationsh |
| | ATT+ | changed link to the color black to convey a positive relationship |
| | ATT2L | changed link to low level of impact |
| | ATT2M | changed link to moderate level of impact |
| | ATT2H | changed link to high level of impact |
| DEL | DEL | deleted the link |
| MOVE | MS | moved a node (which was the same node as the last moved node) |
| | MDn | moved node to the north of the previously moved node |
| | MDne | moved node to the NE of the previously moved node |
| | MDe | moved node to the East of the previously moved node |
| | MDse | moved node to the SE of the previously moved node |
| | MDse | moved node to the South of the previously moved node |
| | MDsw | moved node to the SW of the previously moved node |
| | MDw | moved node to the West of the previously moved node |
| | MDnw | moved node to the NW of the previously moved node |
| COMM | COM | added comment to link to explain how node influences affected node |
| | CREV | revised the existing comment on the given link |

**Codes**

| | |
|---|---|
| 1 | ADD LINK |
| 2 | RELINK CAUSE |
| 3 | RELINK EFFECT |
| 4 | ATTRIBUTE |
| 5 | DELETE LINK |
| 6 | MOVE |

Figure 4. Codes assigned to each action students perform in jMAP while constructing an argument diagram.

**Frequency matrix**

| | ADD LINK | RELINK CAU | RELINK EFFI | ATTRIBUTE | DELETE LINK | MOVE | Replies | No Replies | Givens | % replies | % givens |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ADD LINK | **40** | **8** | **4** | **23** | 1 | **53** | 129 | 2 | 131 | .13 | .13 |
| RELINK CAI | 1 | **1** | 0 | **3** | 0 | **6** | 11 | 0 | 11 | .01 | .01 |
| RELINK EFI | 0 | 0 | 0 | 0 | 0 | 6 | 6 | 0 | 6 | .01 | .01 |
| ATTRIBUTE | **8** | 0 | 0 | **3** | 0 | **19** | 30 | 1 | 31 | .03 | .03 |
| DELETE LII | 3 | 0 | 0 | 0 | **2** | **5** | 10 | 0 | 10 | .01 | .01 |
| MOVE | **79** | **2** | **2** | **2** | 7 | **735** | 827 | 9 | 836 | .82 | .82 |
| | 131 | 11 | 6 | 31 | 10 | 824 | 1013 | 12 | 1025 | | |

**Transitional probability matrix**

| | ADD LINK | RELINK CAUSE | RELINK EFFEC | ATTRIBUTE | DELETE LINK | MOVE | Replies | No Replies | Givens | Reply Rate |
|---|---|---|---|---|---|---|---|---|---|---|
| ADD LINK | **.31** | **.06** | **.03** | **.18** | .01 | **.41** | 129 | 2 | 131 | .98 |
| RELINK CAI | .09 | **.09** | .00 | **.27** | .00 | **.55** | 11 | 0 | 11 | 1.00 |
| RELINK EFI | .00 | .00 | .00 | .00 | .00 | 1.00 | 6 | 0 | 6 | 1.00 |
| ATTRIBUTE | **.27** | .00 | .00 | **.10** | .00 | **.63** | 30 | 1 | 31 | .97 |
| DELETE LII | .30 | .00 | .00 | .00 | **.20** | **.50** | 10 | 0 | 10 | 1.00 |
| MOVE | **.10** | **.00** | **.00** | **.00** | .01 | **.89** | 827 | 9 | 836 | .99 |
| | 131 | 11 | 6 | 31 | 10 | 824 | 1013 | 12 | 1025 | .37 |

**Z-Scores identify probabilities that are higher/lower than expected**

| | ADD LINK | RELINK CAU | RELINK EFF | ATTRIBUTE | DELETE LIN | MOVE | |
|---|---|---|---|---|---|---|---|
| ADD LINK | **6.55** | **6.00** | **3.97** | **10.43** | -0.26 | **-12.56** | 129 |
| RELINK CAI | -0.38 | **2.58** | -0.26 | **4.69** | -0.33 | **-2.29** | 11 |
| RELINK EFI | -0.95 | -0.26 | -0.19 | -0.44 | -0.25 | 1.18 | 6 |
| ATTRIBUTE | **2.28** | -0.58 | -0.43 | **2.24** | -0.56 | **-2.57** | 30 |
| DELETE LII | 1.62 | -0.33 | -0.25 | -0.56 | **6.11** | **-2.56** | 10 |
| MOVE | **-6.76** | **-5.47** | **-3.07** | **-10.98** | -0.96 | **12.98** | 827 |
| | 131 | 11 | 6 | 31 | 10 | 824 | 1013 |

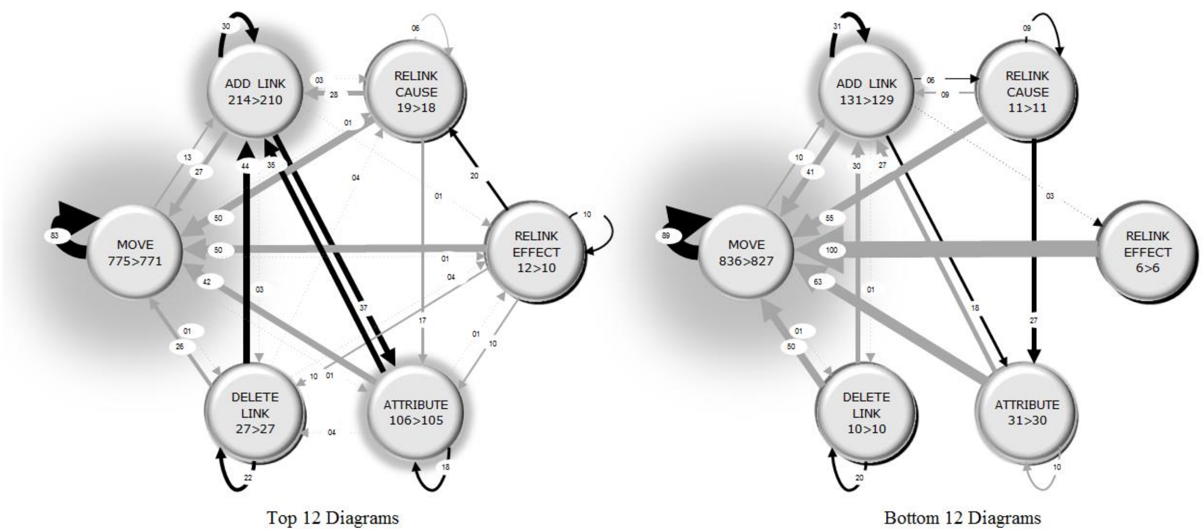
*Note*: The values identified in bold/underline identify action sequences occurring at higher/lower than expected frequency based on the critical *z*-score of ±1.96 at $p < .05$.

Figure 5. Screen shot from DAT showing the frequency, transitional probability, and *z*-score matrices used to reveal sequential patterns in the actions used to create the 12 bottom diagrams.

## Main Findings

The two state diagrams (Figure 6) reveal four action-action sequences that were common in both groups (AddLink→AddLink, MoveNode→MoveNode, and DeleteLink→DeleteLink, AddLink→ChangeLinkAttribute). The first three action sequences suggest for example that students constructed their diagrams using a stage-like sequence by moving multiple premises into position first, then inserting links to connect the premises with links, and then deleting (or correcting) the links between premises. Given that these four processes were observed in both groups, the findings suggest that the use of these four processes neither increases nor decreases the accuracy of students' argument diagrams.

The diagrams also show that top scorers exhibited five unique action sequence patterns. For example, the left diagram shows that when top scorers deleted a link, they were most likely (44%) to follow that action by adding a new link between nodes than bottom scorers. When they specified the attribute of the link, they were most likely (35%) to follow that action by adding a new link. Overall, the differences between the two state diagrams suggest that the following five action sequences (when performed on a consistent basis) can help students construct more accurate diagrams: DeleteLink→AddLink, Attribute→AddLink, Attribute→Attribute, Relink Effect (move the head of the arrow to point to another affected node)→Relink Effect, and Relink Effect→Relink Cause (move tail of arrow to point to another causal node). In particular, the sequence of DeleteLink→AddLink may be an indication of times when students are restructuring their diagrams to undo an error produced when making hasty generalizations (when A→Conclusion and B→Conclusion should in fact be restructured to A→B→Conclusion).




*Note*: Thickness of arrow conveys strength of transitional probability; dark black arrows identify probabilities that are significantly greater than expected based on *z*-score tests ($p < .01$) performed in the DAT software; first and second numerical value displayed in nodes identify the number of times the given action was performed and the number of events that followed the given action; the size of the glow emanating from each node conveys the number of times the action was performed.

Figure 6. State diagrams of processes used to produce the top vs. bottom diagrams.

In contrast, the low-performing students exhibited four unique action sequence patterns (AddLink →RelinkCause (change tail of link to point to a supporting or subordinate premise), AddLink→RelinkEffect (change head of link to point to a superordinate premise), Relink Cause→Relink Cause, and Relink Cause→ChangeLinkAttribute. Given that low scorers exhibited the tendency to perform the RelinkCause→RelinkCause sequence whereas high scorers exhibited the tendency to perform the RelinkEffect→RelinkEffect sequence, these differences suggest that the processes used to: a) create more accurate argument diagrams involved the use of a forward or bottom up approach - systematically examining which major premise (or effect) is supported by a given minor premise (or cause); and b) create the less accurate diagrams involved the use of a backward or top-down approach by progressively examining which minor premise supports a given major or superordinate premise.

To determine if the differences in the observed patterns are statistically significant, a phi-coefficient (ø) will be used to measure the strength of association between group membership and particular pattern (Bakeman & Gottman, 1997). Retroactively recoding and further analysis of the action logs will examine actions (and the actions preceding them) that produced correct versus incorrect links across the argument diagrams of all 17 students.

## Conclusions and Implications

Although the findings are not conclusive, the observed differences in action sequences used by students to construct argument diagrams with high versus low accuracy provide initial insights into the types of processes that can help students create more accurate and achieve deeper understanding of complex arguments. At minimum, the findings and the possible interpretations of the meaning behind the findings provides some insights into the possible methods and approaches that can be used in future research to model and better understand the processes that can be used effectively to analyze and better understand complex arguments.

Future work is needed to: a) replicate this study with larger sample sizes; b) apply multiple approaches to establish validity in the criterion used to assess diagram accuracy, c) refine the precision of the data mining codes in jMAP to fully determine if students are in fact linking premises using a forward or backward approach and are sequentially relinking premises into logical chains to correct for errors produced by making hasty generalizations; d) identify which diagramming actions reflect general reasoning processes that improve argument analysis (areas B and E in Figure 7, respectively); e) integrate the target action sequences directly into the software interface of the diagramming software so that controlled experiments can be conducted to determine cause-effect relationships between target processes and map accuracy; f) determine to what extent particular processes are dependent on students' prior knowledge in order to identify the target processes that can be promoted regardless of students' prior knowledge; and g) examine to what extent the target processes are effective across arguments varying in hierarchical structure/complexity.

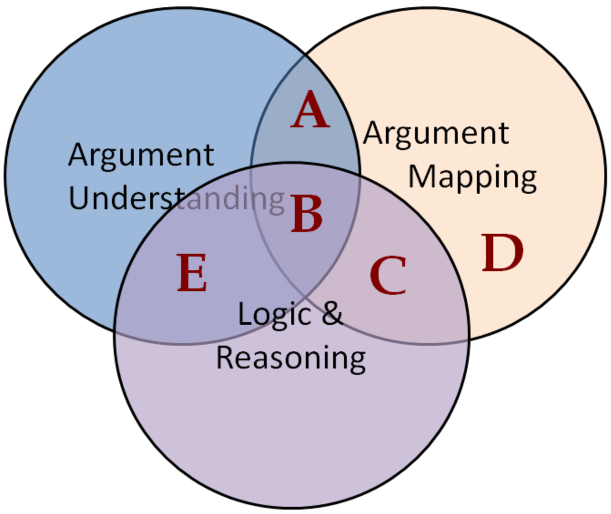



Figure 7. Areas for further research on the relationships between diagramming processes, general reasoning processes, and understanding of complex arguments.

## References

Bakeman, R., & Gottman, J.M. (1997). *Observing Interaction: An Introduction to Sequential Analysis*. Cambridge: Cambridge University Press.

Braak, S. W. van den., Oostendorp, H. van., Prakken, H., & Vreeswijk, G. A. (2006). A critical review of argument visualization tools: Do users become better reasoners? In F. Grasso, R. Kibble, & C. A. Reed (Eds.), *Workshop Notes of the ECAI-2006 Workshop on Computational Models of Natural Argument* (pp. 67-75). Riva del Garda:Italy.

De Neys, W. (2006). Dual processing in reasoning two systems but one reasoner. *Psychological Science, 17*(5), 428-433.

Kuhn, D. (1991). *The Skills of Argument*. Cambridge: Cambridge University Press.

McLean, Cheryl (2005). Evaluating critical thinking skills: Two conceptualizations. *Journal of Distance Education*, *20*(2), 1-20.

Paul, R., & Elder, L. (2001). *The Miniature Guide to Critical Thinking: Concepts & Tools*. Dillon Beach, CA: The Foundation for Critical Thinking.

Ruiz-Primo, M., & Shavelson, R. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, *33*(6), 569-600.Scavarda et (2006).

Scavarda, A.J., Bouzdine-Chameeva, T., Goldstein, S.M., Hays, J.M., & Hill, A. V. (2006). A methodology for constructing collective causal maps. *Decision Sciences Journal*, *37*(2), 263-284.

Sharma, T., Tiwari, N., & Kelbar, D. (2012). Study of difference between forward and backward reasoning. *International Journal of Emerging Technology and Advanced Engineering*, *2*(10), 271-273. Retrieved from http://www.ijetae.com/files/Volume2Issue10/IJETAE_1012_48.pdf

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

van Gelder, T. J. (2007). The rationale for Rationale™. *Law, Probability and Risk*, *6*, 23-42. doi:10.1093/lpr/mgm032.