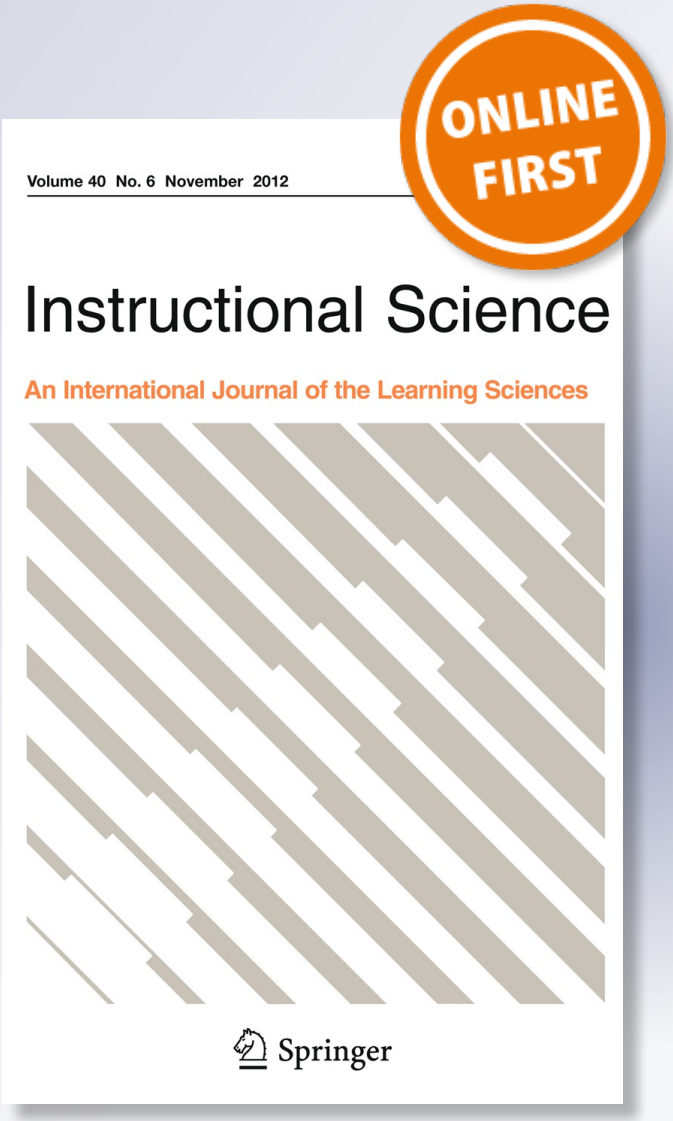
Learning achieved in structured online debates: levels of learning and types of postings

Li Jin & Allan Jeong

Instructional Science
An International Journal of the Learning Sciences

ISSN 0020-4277

Instr Sci
DOI 10.1007/s11251-013-9269-2



Springer

Instr Sci
DOI 10.1007/s11251-013-9269-2

Learning achieved in structured online debates: levels of learning and types of postings

Li Jin · Allan Jeong

Received: 28 September 2011 / Accepted: 4 January 2013

Abstract The purpose of this study was to examine the learning process exhibited in restrained online debates in terms of to what extent each of Bloom's six levels of cognitive learning were exhibited among four types of message (argument, critique, evidence, and explanation). Thirty-three graduate students enrolled in an online entry-level course in distance learning participated in this study. The threaded discussion postings from four structured online debates were analyzed. The results indicated that five of the six levels of cognitive learning (except for knowledge learning) were achieved to a certain extent among the four types of postings. Chi square tests indicated that higher levels of learning were most likely to be exhibited in critique and argument postings. Message–response exchanges ending with critiques or starting with argument messages were most likely to elicit higher levels responses. The results were discussed in light of better understanding of student learning and implications for instructional designs of discussion topics and restraints.

Keywords Structured online debate · Constraint-based online debate · Levels of learning · Learning outcome · Online argumentation

Introduction

Structured online debates have been examined as a way to enhance cognitive learning (Jeong and Juong 2007; Moore and Marra 2005). Structured online debates, also referred to as constraint-based argumentation (Cho and Jonassen 2002) or scripted online argumentation (Stegmann et al. 2007; Weinberger et al. 2005), are instructional activities in online

L. Jin (✉)
Department of Educational Psychology and Learning Systems, The Florida State University, 184-06 Moore Dr., Tallahassee, FL 32310, USA
e-mail: ljin@fsu.edu

A. Jeong
Department of Educational Psychology and Learning Systems, The Florida State University, 3205 E Stone Building, Florida State University, Tallahassee, FL 32306, USA

Published online: 11 January 2013

discussions that specify the types of posting students can contribute to the discussions, such as claim, warrant, rebuttal, and backing, which are developed based on certain argumentation models. Argumentation is an essential process of formulating ideas and solving problems (Kuhn 1991) and is advocated as an effective way to promote deeper understanding of a subject (Baker 2003). The use of constraints is to engage students in the process of constructing effective argumentations and consequently achieve high level of knowledge learning (Schellens and Valcke 2005), which involves cognitive processes of analysis, synthesis, and evaluation.

When engaged in argumentation, two types of learning outcomes are expected: one is the acquisition of argumentation skills, and the other is the acquisition of the subject knowledge based on which the arguments are constructed. The quality of the argumentation is hypothesized to be positively related to the acquisition of subject knowledge (Stegmann et al. 2007). However, the research on this topic is limited and inconclusive. This study investigates the impact on subject knowledge acquisition and aims to provide reference to instructors on whether an restraint-based argumentation approach in online discussions should be adopted when the instructional goal is other than improving argumentation skills.

Inconsistent results have been reported in the few studies that investigated the effectiveness of structured debates in enhancing subject knowledge learning. In the Stegmann et al. (2007) study, the discussants were provided with three scripts (claim, ground, and qualification) on construction of each posting and three scripts on the types of posting they should contribute to the debate (argumentation, counter-argumentation, and integration). The average gain of pre-test and post-test showed that the use of restraints did not affect the domain-specific knowledge acquisition. In the study of Cho and Jonassen (2002), students in the treatment group were expected to contribute to the debates on one of the five types of messages: claim, ground, warrant, backing, and rebuttal, which were the major components of Toulmin's (1958) argumentation model. Compared to the students who used conventional threaded discussion design, no difference was found in their performance in problem-solving tasks. Based on the same argumentation model of Toulmin's, Moore and Marra (2005) developed four restraints, thesis, evidence, assumption, and synthesis for the online discussants. These authors used a content analysis approach to the assessment of student knowledge construction. They used the Interaction Analysis Model (Gunawardena et al. 1997) to analyze student postings. This model posited that learners would move through five phases when knowledge was being constructed. The higher phases a learner reached the better learning he achieved. They found significant difference in the number of postings coded as higher phases between the constraint-based and the non-constrained discussion groups, which indicated a negative impact of the use of restraints. It seems that the overall result from the empirical studies on the effect of restraint-based online discussion on knowledge acquisition is not as promising as it has been expected.

The reason previous research has failed to provide supporting evidence on the use of restraints in online discussion could be that the learning process, which the argumentation was meant to enhance, was ignored. In order to arrive at strong arguments, the students have to search for related ideas, consider multiple perspectives, develop or select a solution, and justify proposed solutions with supporting evidence (Kobbe et al. 2007; Cho and Jonassen 2002). The use of restraints is to engage the students in these processes and eliminate the behaviors on off-task discussions, simple agree–disagree exchanges, or superficially sharing of existing knowledge and experience (Jeong and Juong 2007). As a result, the students are expected to construct meaningful interactions that may help them

process the knowledge at high cognitive levels. Unfortunately, scant research has been conducted on this topic. In order to enrich the literature, we investigate the cognitive presence exhibited instructured online debates in this study.

The students were provided with four restraints in this study: argument, critique, explanation, and evidence. Argument was used to make claims of the proposition, convey reasons to support a given position on a given issue. Critique was to challenge or critique given propositions, evidence or clarifications. Data, theoretical supports, or definitions should be included to justify the critique. Evidences were examples, research studies, personal observations and experiences that would support the position held by the individual. Explanation was the elaboration or clarification of the opinion. These four types of messages are the core elements of Toulmin's (1958) model and represented the conflict in interactions and the inquiry triggered by the conflicts (Jeong and Juong 2007). Therefore, these constraints should be able to direct the students in the process of identifying issues, constructing ideas, confronting discrepancies, and justifying propositions.

In order to investigate the discussion processes in terms of the cognitive levels being achieved, we used Bloom's (1956) taxonomy of cognitive learning as the assessment criteria. There are six levels of cognitive learning outcomes defined in this hierarchy: knowledge, comprehension, application, analysis, synthesis and evaluation, sequenced from low to high in cognitive level. In knowledge and comprehension learning, learners show recall and understanding of materials, processes, and ideas. Application is the use of learned materials (rules, laws, methods, theories, etc.) in new contexts; analysis is the breakdown of communication into its constituent elements or parts such that the structures of the idea or the relation embedded are made explicit. Synthesis is the ability to reorganize parts together to form a new whole, which involves the process of working with pieces from different sources to constitute a new pattern or new structure. Evaluation is process of making purposefully judgment and justifying it with supports, evidence, data, and reasons. Analysis, synthesis, and evaluation are often referred to as high levels of learning.

Constrained by the types of message that could be posted in online debates, students were engaged in the argumentations that require high levels of cognitive processing, such as analyzing, evaluating, and constructing (Cerbin 1988). Students must be able to identify the key points of a statement in order to distinguish between claims, grounds, or rebuttals. In organizing lines of statements from both sides of a proposition, the students must be able to construct new claims based on multiple ideas. They also needed to apply a set of criteria for judging whether claims are well-supported or warrants are adequately backed. As a result, the students were expected to engage in the cognitive processes of higher levels of cognitive learning outcomes in Bloom's (1956) taxonomy: analysis, synthesis, and evaluation (Valcke et al. 2009).

The use of Bloom's taxonomy in the analysis may enable us to learn how the students have processed the content knowledge in structured online debates, based on which we can make inference about how structured online debates may have facilitated learning. If we can identify the relationships between types of constrains and levels of learning, we may be able to predict which types of messages are more likely to elicit higher level cognition. As a result, we can better understand the dynamic association between learning process and constraints. This type of knowledge in turn may shed lights on how to improve the design and implementation of constraint-based argumentation.

The purpose of this study is to investigate the learning achieved in structured online debates in terms of the levels of learning achieved in each type of postings. Specifically,

this study aims to determine to what extent each level of cognitive learning has been exhibited within each of the four types of postings used in online debates (argument, critique, explain, evidence). The second purpose of this study is to identify which type(s) of message is more likely to exhibit a particular level of cognitive learning in Bloom's taxonomy. A third purpose of this study is to identify what types of exchange pairs (e.g. argument → critique, critique → explanation) are most likely to elicit or trigger responses that exhibit higher levels of learning (e.g., analysis, synthesis, and evaluation). In summary, the questions examined in this study are:

1. To what extent did each type of posting exhibit each level of cognitive learning in Bloom's taxonomy?
2. Which type of posting was more likely to exhibit each particular level of learning in Bloom's taxonomy?
3. What exchange pairs tended to exhibit higher levels of learning in Bloom's Taxonomy?

Methods

Participants

Participants were 33 graduate students enrolled in an online introductory course of Distance Education in a large southeastern public university in the United States. Twenty-two were female and eleven were male. All of them were from the program of Instructional Systems. Most of them were distance learners who had a part-time or full-time job. All the students consented to participate in this study. Participation in the online debates was voluntary and would not be counted toward the course grades.

Procedures

Students in the online course participated in a total of 4 weekly online debates on given topics. The topics for these four debates were: media selection and learning, the use of synchronous tools in online courses, cost issue of educational technology, and printing versus electronic materials for education. The instructor posted the discussion topics to the discussion board (on Blackboard) at the beginning of each week but did not actively participate in the debates. The purpose of the debates was to engage students in sharing diverse perspectives of the topics, reflecting on their view points, reasoning based on evidence, and supporting or disputing ideas with data. The actual number of participants in each of the online debates varied from 12 to 15 students.

For each debate, the participants were randomly assigned to one of the two teams: supporting team and opposing team. They were required to post one of the four types of messages when they support or refute the point: argument, evidence, critique, and explanation. Descriptions of what each constraint meant were provided to participants at the beginning of the course. Participants were required to insert tags (ARG, EVID, BUT, EXPL) into the subject headings of each posting to identify the types of message they posted. Tags had to include a positive (+) or negative (−) sign (e.g. +ARG, −BUT) to identify the authors' team membership (supporting team or opposing team). See Fig. 1 for a screen shot of the threads and tags in a structured online debate.

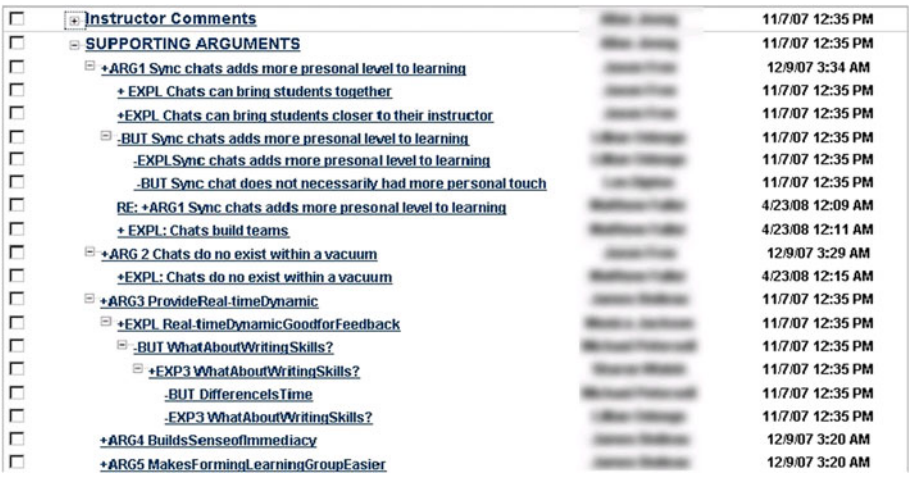


Fig. 1 Screen shot of threads and tags in a structured online debate

Measures

Level of cognitive learning

Levels of cognitive learning exhibited in the structured debates were assessed using Bloom's (1958) taxonomy of cognitive learning. The six levels of learning from low to high were: knowledge, comprehension, application, analysis, synthesis, and evaluation. Each posting in the four debates was coded into one of the six levels of learning. A message was assigned as the highest level it exhibited when it indicated more than one level of learning. See Table 1 for the description of the categories and examples of the postings. A second coder coded one debate and a comparison of the coding results between the primary and second coder produced an inter-rater agreement rate of 81.4 %.

Type of posting

Each of the student postings was coded as one of the four restraints: argument, evidence, critique, and explanation. Types of posting were coded based on students' self-labeling of each message. Whether a student label matched the content of the posting was checked by the researchers before conducting the analysis. Mislabeling happened when a student labeled his/her posting with a wrong type of constraints (e.g., labeling an argument as a critique), or did not label the posting with one of the four constraints. Mislabeling rate was the percentage of postings that were labeled incorrectly over the total number of postings. The overall mislabeling rate of the four debates was 13.3 %. The mislabeling rate for each of the four debates was 9.3 % (7 out of 75), 17.7 % (14 out of 79), 19 % (15 out of 79), and 5.9 % (4 out of 69), respectively. Over half of the mis-labels (21 out of the 40 cases) were between explanation and evidence. This indicated that the students might have paid less attention on the difference between evidence (to support a claim) and explanation (to clarify a claim). Another major reason that a message was mislabeled was because when the students used the "reply" feature in Blackboard (8 cases), the subject line of the previous message was automatically placed as the heading of the new message and some

Table 1 Description of the categories and examples of student postings

Learning	Description	Example
Knowledge	Recall of factors, methods, processes, patterns, etc	Educational technology refers to the equipments, media, and ways we present the instructions
Comprehension	The understanding of the materials or ideas being communicated	The data suggests as Clark states that, "the delivery system affects no inherent difference on achievement." So although I can understand the crux of your explanation, I would like to disagree with your conclusion
Application	The use of abstractions in particular concrete situations	…sometimes instructors use these technologies when they aren't really necessary. For example, would you really need to see leaves changing color for fall on a computer screen…or could you just go outside?
Analysis	The breakdown of an idea into its constituent elements and make their relations explicit	(Because)Students tend to give increased effort and attention to media (such as TV) that are new to them. It's not the media that yields achievement gains, it's the increased effort and persistence
Synthesis	Working with pieces, parts, elements, etc.…and arranging and combining them into new patterns or structures	It probably does not matter in cases where the objectives to be met are simple and clear. But what about cases where there are cognitive or psychomotor skills to be learned that are more complex, such as learning to perform laser eye surgery. Something like this would be difficult to learn from a teacher who is just showing pictures in a book and reading "how to" steps
Evaluation	Judgments about the merits of Materials and methods for given purposes	As a former Elementary Teacher I can tell you that how students think they learn best…and how they actually learn best are sometimes two different things. If this study was based on students' perceptions, is it really a genuine assessment of learning?

students forgot to change it. The types of posting were corrected to be included in the next step analysis.

Data analysis

In order to test whether one (or some) of the categories of the outcome (level of learning and type of posting) is more likely to occur than other categories, a χ^2 goodness of fit test was conducted to determine whether the categories under each measure (level of learning and type of posting) were equally distributed. The assumption for this test was the frequency for observing any types of postings (or level of learning) in the data was equal. The χ^2 test would be significant if the observed distribution was not statistically equal across the categories. Because there were more than two categories in each measure, the test results would not tell which frequency was significantly lower or higher than that of other

Table 2 Descriptive results of the four debates

Debate	No. of participants	Types of posting				
		Argument	Critique	Evidence	Explanation	Total
1	15	15	30	17	10	72
2	15	21	20	22	15	78
3	15	26	38	8	7	79
4	12	26	26	8	9	69
Total		88	114	55	41	298
M (SD)		22.0 (5.2)	28.5 (7.5)	13.8 (6.9)	10.3 (3.4)	74.5 (4.8)

categories. We would examine the data and choose the highest value as the mostly observed category.

Results

As shown in Table 2, a total of 298 postings was coded and analyzed. Among the four types of postings, critique was the mostly observed (114, $M = 28.5$, SD $= 7.5$), argument was the second (88, $M = 22.0$, SD $= 5.2$). Evidence postings totaled 55 ($M = 13.8$, SD $= 6.9$) and the number of explanation posting was 41 ($M = 10.3$, SD $= 3.4$).

RQ 1 To what extent did each type of posting exhibit each level of learning in Bloom's taxonomy?

Levels of learning exhibited in each type of posting are reported in Table 3. No knowledge level learning has been observed (this probably was because a message was coded as the highest level of cognitive learning it exhibited). Therefore, knowledge category was eliminated from the subsequent analyses.

Arguments

Since there was no evaluation level cognition observed in argument messages, a χ^2 test was conducted for the comprehension, application, analysis, and synthesis levels. The result, $\chi^2(3, N = 88) = 9.55$, $p = .03$ indicated that there was significant difference among the occurrences of the four levels of learning. By examining the data, we concluded that

Table 3 Number counts of messages exhibiting each level of learning in each type of postings

Levels of learning	Types of postings				
	Argument	Critique	Evidence	Explanation	Total
Knowledge	0	0	0	0	0
Comprehension	28	17	24	13	82
Application	11	18	20	9	58
Analysis	29	57	9	16	111
Synthesis	20	14	2	1	37
Evaluation	0	8	0	2	10

comprehension (31.8 %) and analysis (33.0 %) levels were mostly achieved in argument postings.

Critique

The χ^2 test (4, $N = 114$) = 66.8, $p < .05$) suggested that the analysis learning was mostly achieved in critique messages.

Evidence

No evaluation was observed in evidence messages. The χ^2 test was conducted among the four levels that were observed. The result $\chi^2(3, N = 55) = 22.16$, $p < .05$ indicated that comprehension learning was mostly achieved in evidence messages.

Explanation

The χ^2 test was (4, $N = 41$) = 21.37, $p < .05$. By examining the data, we concluded that analysis was mostly achieved in this type of postings.

RQ 2 Which type of posting was more likely to exhibit each particular level of learning?

Table 4 showed the distribution of each level of learning achieved in each type of message. A χ^2 test comparing the observed and expected frequencies of each type of posting in specific levels of learning indicated which type of posting was more likely to exhibit a specific level of learning. We conducted χ^2 tests for the levels that had non-zero entries: comprehension, application, analysis, and synthesis. For evaluation level, we conducted the test between two non-zero entries: critique and explanation.

The χ^2 tests showed no significant difference for comprehension learning, $\chi^2(3, N = 82) = 6.68$, $p = .08$, or application learning, $\chi^2(3, N = 58) = 5.86$, $p = .12$, among the four types of message. However, significant difference was found in the higher levels of learning outcome, $\chi^2(3, N = 111) = 48.53$, $p < .05$, for analysis, $\chi^2(3, N = 37) = 29.97$, $p < .05$, for synthesis, and $\chi^2(1, N = 10) = 17.20$, $p < .05$, for evaluation. By reviewing the data, we conclude that critiques were most likely to elicit analysis learning, arguments and critiques were most likely to elicit synthesis learning, and critiques were most likely to elicit evaluation learning.

RQ 3 What exchange pairs tend to elicit responses that exhibit higher levels of learning?

Table 4 Overall distribution of each level of learning in each type of posting (in percentage)

Types of posting	Levels of learning					
	Knowledge	Comprehension	Application	Analysis[a]	Synthesis[a]	Evaluation[a]
Argument	0	34.1	19.0	26.1	54.1	0
Critique	0	20.7	31.0	51.4	37.8	80.0
Evidence	0	29.3	34.5	8.1	5.4	0
Explanation	0	15.9	15.5	14.4	2.7	20.0

[a] χ^2 test for this column is significant at .05

Table 5 Number counts of exchange pairs exhibiting higher levels of learning in the ending message

Exchange pair	Higher levels of learning			
	Analysis	Synthesis	Evaluation	Total
Argument–argument	7	4	0	11
Critique–argument	1	1	0	2
Evidence–argument	0	2	0	2
Explanation–argument	0	0	0	0
Subtotal	8	7	0	15
Argument–critique	27	11	5	43
Critique–critique	17	2	3	22
Evidence–critique	11	1	0	12
Explanation–critique	2	0	0	2
Subtotal	57	14	8	79
Argument–evidence	10	2	0	12
Critique–evidence	2	0	0	2
Evidence–evidence	0	0	0	0
Explanation–evidence	0	0	0	0
Subtotal	12	2	0	14
Argument–explanation	13	0	0	13
Critique–explanation	4	1	1	6
Evidence–explanation	0	0	0	0
Explanation–explanation	0	0	1	1
Subtotal	17	1	2	20
Total	94	24	10	128

An exchange pair consisted of the initial posting and the posting that responded to it. An initial argument posting and the critique posting that replied to the argument was an exchange pair of argument →critique. We examined the level of learning exhibited in the responding message (the ending message in the exchange pair) and reported the number counts of the exchange pairs that exhibited higher levels of learning, analysis, synthesis, and evaluation, in the ending message (see Table 5).

A χ^2 test indicated that the frequency of eliciting an analysis response in the responding message among the observed exchange pairs was not equal $\chi^2(9, N = 94) = 63.66, p < .05$. By examining the data, we concluded that argument–critique exchange pairs were most likely to exhibit analysis learning in the responding messages. The χ^2 test for the equal distribution of exchanges elicited synthesis responses was significant, $\chi^2(7, N = 24) = 26.67, p < .05$. It seemed that argument–critique exchanges were most likely to exhibit synthesis learning in the responses than other exchange patterns. Significant difference was found for the frequencies of the exchange pairs that ending with evaluation learning, $\chi^2(3, N = 10) = 11.2$, $p = .01$. It seemed that argument–critique pair was most likely to elicit evaluation learning in the responding message.

Post-hoc analysis

Post-hoc analysis was conducted on the emerging pattern from the data. As shown in Table 6, the exchange pairs that started with argument messages were most likely to elicit

Table 6 Number counts of exchange pairs eliciting higher levels of learning in the ending message

Exchange pair started with	Higher levels of learning		
	Analysis	Synthesis	Evaluation
Argument	57	17	5
Critique	24	4	4
Evidence	11	3	0
Explanation	2	0	1

responses that exhibited higher levels of learning. At the same time, an explanation message was the least likely to elicit a response that exhibited higher levels of learning.

Discussion and conclusions

The purpose of this study was to investigate the cognitive process the students engaged in during structured online debates. Four types of restraints were used in this study and Bloom's taxonomy of cognitive learning outcome was used to analyze student postings. About 15.8 % of the postings in the debates were at the synthesis and evaluation level. Compared to the study conducted by Garrison et al. (2001), where they reported $\sim$17 % of the student postings showed construction of solutions or critical assessment of ideas, the students used structured online debates in this study achieved a similar level of learning compared to the students used un-restrained online discussion in the Garrison et al. (2001) study. However, compared to the results from Gunawardena et al. (1997), where $<$7 % of the pre-service teachers' discussion were at higher level knowledge processing, the structured debates in our study seemed have enhanced student learning.

No knowledge level learning has been observed in any types of constraints. Students did not present only fact knowledge in their postings but indicated applying their knowledge in the context of the debate topics, or reached analyzing and synthesizing of the related materials. This probably is because the nature of the debates. Students are required to construct claims to support or argue against given propositions in the debates. Although students must have sufficient topic knowledge (Cerbin 1988), they need to make the decision on what knowledge is relevant to the ongoing discussion and make the connections in their postings. This process elevates the students' cognitive learning to at least comprehension level.

As to research question 2, the results shows that students are more likely to exhibit high levels of learning in critiques and arguments and less likely to engage in high level processing in constructing evidence postings. These results indicate the likely effects of constraints on level of cognitive processing. First, argument and critique constraints are very likely to exhibit higher levels of learning because the specifications of these two constraints require students to include reasons or clarification to support or justify their propositions in the same posting. Second, evidence constraint is less likely to engage students in high levels of cognitive processing because the constraints seems to request for factual information that are judged as relevant to the issues being discussed. Based on these findings, it seemed that instructors could facilitate student learning at different level by manipulating the debate restraints. If the purpose of the online discussion was to help students understanding the subject knowledge and make appropriate applications (which are common for entry level courses, or to review pre-knowledge at the beginning of an instruction), the prompts that requesting students to explain or providing evidence would

work. When the difficulty level proceeds, and much new knowledge needs to be constructed rather than memorized, and the conflicting ideas or opinions are likely to occur among the students, restraints directing the students to argue for their proposition and critique other's statements may be of much help.

For research question 3, the results show that exchange pairs that ended with critiques were most likely to demonstrate high levels of learning. Further investigation of the patterns of exchange pairs and learning outcomes indicates that argument is most likely to elicit higher level responses (post hoc analysis). This may be due to the fact that an argument presented a proposition that lays the ground for critique, and when students construct critiques they may have conducted analysis of relevant ideas and opinions and their learning have been elevated to levels that above analysis. Based on this finding, we suggest that online instructors pay attention to the frequency of argument and critique postings. More argument postings are likely indicating diverse opinions and active participation in the debates. If the number of critique postings is low in a discussion, it is likely that the students do not engage in the debates at a meaning constructing level or do not have sufficient information on the discussion topics. Thus, certain interventions should be applied to encourage them to reflect, critique, and integrate ideas. Alternatively, the instructors may consider using different topics for the debates.

In conclusion, the results of this study show that the use of constraints in online debates engaged students in the processes of analyzing ideas, identifying relations, making decisions, and constructing new knowledge and consequently facilitated student learning. However, the results of this study did not contradict to previous findings (Cho and Jonassen 2002; Moore and Marra 2005) of no-improvement in student test or performance. We consent that although the use of structured online debates may engage the students in a meaningful inquiry process, this does not necessarily lead to improved performance in consequent tests because participation in the debates may not sufficient for the students to master the knowledge and skills to a certain extent that would assure their better performance in the follow-up tasks.

There are several limitations to this study that restrict the generalization of its results. First, the operational definitions of evidence and explanation restraints may not be very clearly distinguishable and as a result close to half of the mislabeled messages were between these two restraints. We recommend that future applications or studies using similar concepts should provide distinctive descriptors. One other limitation is related to the participants and the subject matter. The sample consisted of graduate students enrolled in an online course in the field of education. Students at a different level, such as undergraduate programs, may exhibit different patterns in online debates. At the same time, instructors may need to be cautious about the application of the structured debate in other subject fields such as math or science because the patterns of interaction may be very different. The authors of this study also recommend that instructors develop their own restraints to facilitate student debates in their courses. The use of the course-specific discussion restraints may better engage the students in the learning processes that are desired for the purpose of the course.

References

Baker, M. J. (2003). Computer-mediated argumentative interactions for the collaboration of scientific notions. In J. Anderiessen, M. J. Baker & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 47–78). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Bloom, B. (1956). Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners. In B. Bloom (Ed.), *Handbook 1: Cognitive domain*. New York: Longmans.

Cerbin, B. (1988). *The nature and development of informal reasoning skills in college students*. ERIC document No. ED 298 805.

Cho, K., & Jonassen, D. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development, 50*(3), 5–22.

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education, 15*(1), 7–23.

Gunawardena, C., Lowe, D., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computer Research, 17*(4), 397–431.

Jeong, A., & Juong, S. (2007). Scaffolding collaborative argumentation in asynchronous discussions with message constraints and message labels. *Computers and Education, 48*, 427–445.

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hamalaninen, R., & Fisher, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*(2–3), 211–224.

Kuhn, D. (1991). *The skills of argument*. Cambridge, England: Cambridge University Press.

Moore, J., & Marra, R. (2005). A comparative analysis of online discussion participation protocols. *Journal of Research on Technology in Education, 38*(2), 191–212.

Schellens, T., & Valcke, M. (2005). Collaborative learning in asynchronous discussion groups: What about the impact on cognitive processing? *Computers in Human Behavior, 21*(6), 957–976.

Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *Computer-Supported Collaborative Learning, 2*, 421–447.

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

Valcke, M., de Wever, B., Zhu, C., & Deed, C. (2009). Supporting active cognitive processing in collaborative groups: The potential of Bloom's taxonomy as a labeling tool. *Internet and Higher Education, 12*, 165–172.

Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science, 33*(1), 1–30.