
November – 2005

The Effects of Linguistic Qualifiers and Intensifiers on Group Interaction and Performance in Computer-Supported Collaborative Argumentation

Allan C. Jeong
Florida State University

Abstract

This study tested the effects of linguistic qualifiers and intensifiers on the number and types of replies elicited per argument and per challenge posted in online debates. To facilitate collaborative argumentation, thirty-two students (22 females, 10 males) enrolled in a graduate-level online course classified and labeled their messages as arguments, challenges, supporting evidence, or explanations prior to posting each message. The findings showed that qualified arguments elicited 41 percent fewer replies (effect size = -.64), and the reduction in replies was greatest when qualified arguments were presented by females than males. Challenges without qualifiers, however, did not elicit more replies than challenges with qualifiers. These findings suggest that qualifiers were used to hedge arguments, and such behaviors should be discouraged during initial stages of identifying arguments (more so in all-female than in all-male groups) in order to elicit more diverse and more opposing viewpoints needed to thoroughly and critically analyze arguments.

Keywords: Computer-mediated communication; CMC; communication style; group interaction patterns; interaction analysis; computer-supported collaborative learning; CSCL; collaborative argumentation

Introduction

Computer-mediated communication (CMC) is widely used to support student interaction in order to facilitate higher order learning through critical discussion. Collaborative argumentation is one activity used to foster critical discussion (Johnson and Johnson, 1992) in both face-to-face and online environments. Argumentation involves the process of building arguments to support a position, considering and weighing evidence and counter-evidence, and testing out uncertainties to extract meaning, achieve understanding (McAlister, 2003), and examine complex problems (Cho and Jonassen, 2002). Computer-supported collaborative argumentation (CSCA) provides students the opportunity to practice argumentation through writing and discussion simultaneously while communicating with text-based communication tools (Baker, 1999).

Various strategies have been developed to support collaborative argumentation where constraints are imposed on the types of messages students can post to a discussion. For example, Jeong and Juong (*in press*) presented to students a fixed set of message categories (arguments, challenges, supporting evidence, explanations) and required students to classify and label each message by

inserting a tag corresponding to a given message category in the subject headings of each message prior to posting them to threaded discussions in *Blackboard*, a course management system. Using a more formalized approach, Jonassen and Remidez (2002) developed a threaded discussion tool called *ShadowPDforum* where the message constraints are embedded and built into the computer interface so that students are required to select (from a menu of options) and classify the function of each message before messages are posted to discussions. This particular approach has been implemented in other asynchronous discussion environments like *ACT* (Duffy, Dueber and Hawley, 1998; Sloffer, Dueber and Duffy, 1999), *FLE3* (Leinonen, Virtanen, and Hakkarainen, 2002), *NegotiationTool* (Beers, Boshuizen, and Kirschner, 2004), and also in synchronous internet chat tools like *AcademicTalk* (McAlister, 2003).

Few if any studies, however, provide conclusive evidence to show that message constraints improve students' performance in collaborative argumentation and learning outcomes. Message constraints (or "social scripts") have been found to elicit more replies that elaborate on previous ideas, and produce greater gains in individual acquisition of knowledge (Weinberger, Ertl, Fischer, and Mandl, 2005). In another study, message constraints generated fewer unsupported claims and achieved greater knowledge of the argumentation process (Stegmann, Weinberger, Fischer, and Mandl, 2004). No differences were found, however, in individual knowledge acquisition, students' ability to apply relevant information and specific domain content to arguments, and ability to converge towards a shared consensus. Furthermore, message constraints were found to inhibit collaborative argumentation – producing fewer challenges per argument than argumentation without message constraints (Jeong and Juong, 2005).

These mixed findings suggest that students may require additional forms of guidance beyond what is offered with the use of message constraints. Message constraints provides guidance on "what" types of messages to contribute to discussions, but provides no guidance on "how" best to present one's ideas in ways that foster rather than inhibit critical discussion. Given the contentious nature of argumentation, managing the exchange of opposing viewpoints can be challenging in CMC because many students (in one case, 50 percent or more) prefer not to share ideas on controversial topics in CMC (Austin, 1997) and because non-verbal cues are absent in online discussions (Walther, 1992). Anywhere from 50-70 percent of face-to-face communication is conducted through non-verbal cues (Mehrabian, 1968). Nonverbal cues like crossing of arms, rigid posture, hesitations, and averting eye contact are useful for determining how best to manage confrontations. In addition, vocal pleasantness, physical proximity, and facial expressiveness have been found to be positively associated with judgments of communicator competence and persuasiveness (Burgoon, Birk, and Pfau, 1990). To compensate for the absence of non-verbal cues in CMC, students may need additional guidance on what linguistic forms to use or not to use when presenting arguments in order to foster both meaningful and critical exchanges between discussion participants.

Linguistic forms that are likely to play a role in how students engage in argumentation are linguistic qualifiers and intensifiers. Previous research on these two linguistic forms has been examined in online group discussions (Blum, 1999; Fahy, 2002a, 2002b, 2003; Herring, 1993, 1996; Savicki, Kelly, and Ammon, 2002). Previous studies show that females use more qualifiers than males, and that males use more intensifiers than females (Fahy, 2002a, 2002b). Fahy (2002b) found that females produced 57 percent of the most commonly used qualifiers (e.g., but, if, may, I think, often, probably, though) in instructor-moderated online discussions. The largest difference was in the use of "I think", where 68 percent of the total uses were by females. In contrast, males produced 61 percent of the most commonly used intensifiers (e.g., very, only, every, never, always), with males using "very" almost twice as often as females. Furthermore, females used qualifiers 3.6 times more often than intensifiers, while men used qualifiers 1.7 times more often than intensifiers. Previous studies also show that participants that use qualifiers tend to be

perceived as less persuasive and less credible (Hosman, 1989), particularly more so for females than males (Bradley, 1981). All together, these findings suggest that when qualifiers are used to present arguments, such arguments may be more likely to elicit replies, or more specifically, elicit challenges that question the merits of the argument than those presented without qualifiers. Yet at the same time, qualifiers can serve as hedges to deflect responses from potential challengers.

At this time, no reported studies have examined how students, male or female, in online discussions respond to other participants' messages when the ideas are presented with linguistic qualifiers and intensifiers. Studies are needed to examine: (a) how qualifiers and intensifiers, when used to present arguments and challenges, affect the number of elicited replies; and (b) to what extent do they elicit the types of replies most likely to increase the level of discussion and critical analysis of arguments. These types of questions must be examined in order to understand the strategic value of using various linguistic forms to encourage interaction and engage participants in the processes of verifying (e.g., argument → challenge → evidence) and justifying (e.g., argument → challenge → explain) arguments to improve collaborative work, decision-making, and problem solving in CMC.

Theoretical Assumptions

The nature of the collaborative task, the research questions, and methods used in this study are grounded under the assumptions of the dialogic theory of language (Bakhtin, 1981; Koschmann, 1999). The theory's main assumption is that social meaning is re-negotiated and constructed as a direct result of conflict produced in social interactions, and that conflict is the primary force that drives critical inquiry and dialog. The second assumption is that *conflict* is produced not by the utterance itself, but by the juxtaposition of inter-locking pairs of utterances. As a result, the need to explain, justify, and understand is felt and acted upon only when conflicts or errors are brought to attention (Baker, 1999). Supporting these assumptions are the findings from extensive research on collaborative learning in the face-to-face classroom (Johnson and Johnson, 1992; Wiley and Voss, 1999) and some recent research in CMC (Jeong, 2004b; Lemus, Seibold, Flanagin, and Metzger, 2004) that show conflict (produced by responses that challenge arguments) and the consideration of both sides of an issue is what drives inquiry, reflection, articulation of individual viewpoints and assumptions, and deeper understanding.

Purpose

This study explored how linguistic qualifiers and intensifiers affect the way messages and replies are exchanged when students engage in collaborative argumentation in asynchronous threaded discussions. This study examined four questions:

1. Does linguistic form affect the mean number of replies elicited by arguments and do the differences vary by the gender of the participant posting the argument?
2. Does linguistic form affect the number of replies elicited by challenges and do the differences vary by the gender of the participant posting the challenge?
3. Does the type of linguistic form used to present an argument produce different response patterns, and to what extent do the observed patterns lead to higher levels of critical analysis?
4. Does the type of linguistic form used to present a challenge produce different response patterns, and to what extent do the observed patterns lead to higher levels of critical analysis?

Method

Participants

The participants were graduate students ($n = 32$) from a major university in the Southeast region of the United States, with ages ranging from 20 to 50 years old. Participants were enrolled in a 16-week online graduate introductory course on distance education. Seventeen of these participants (11 females, 6 males) were enrolled in the course during the fall term. The remaining 15 participants were enrolled in the same course in the following term (11 females, 4 males).

Debate procedures

The students in the fall term participated in five debates, and students in the spring term participated in three debates. In both courses, students used threaded discussion forums in *Blackboard*, a Web-based course management system. Furthermore, the online debates in both courses were identically structured. Student participation in the debates and other discussions throughout the course contributed to 20 percent of the course grade. For each debate, students were required to post at least four messages. Prior to each debate, students were randomly assigned to one of two teams (balanced by gender) to either support or oppose a given position. Finally, students were required to vote on the team that presented the strongest arguments following each debate. In all the debates, the instructor did not participate in the debates, but on rare occasions, the instructor posted messages to ensure that students followed the rules and protocols.

In both iterations of the course, the total number of students and the male-to-female ratio were quite similar. The only notable difference between the two iterations is that the number of debates in the spring term was reduced from five debates to three debates to respond to students' who felt that five debates was too many within a single course. As a result, the students in the spring term participated in two fewer debates. The three remaining debates used in both iterations addressed the same topics and issues. The purpose of each debate was to critically examine design issues, concepts and principles in distance learning examined in the course. For example, students debated the following claims: "The Dick and Carey ISD model is an effective model for designing materials for online courses," "The role of the instructor should change when teaching at a distance," and "Type of media does not make any significant contribution to student learning."

Online debate messages and message labels

Students were presented a list of four message categories (see Figure 1) during the debates to encourage students to support and refute arguments with supporting evidence, explanations, and critiques. Based on Toulmin's (1958) model of argumentation, the response categories and their definitions were presented to students prior to each debate. Each student was required to classify each posted message by category by inserting the corresponding label into the subject headings of each message, and restrict the content of their messages to address one, and only one, category at a time. The investigator occasionally checked the message labels to determine if students were appropriately labeling their messages according to the described procedures. No participation points were awarded for a debate when students failed to follow these procedures. Students were able to return to previous messages to correct any errors in the message labels.

Figure 1. Example instructions on how to label messages during the online debates

Symbol	Description of symbol
+	Identifies a message posted by a student assigned to the team <u>supporting</u> the given claim/statement
-	Identifies a message posted by a student assigned to the team <u>opposing</u> the given claim/statement
ARG#	Identifies a message that presents <u>one and only one</u> argument or reason for using or not using chats (instead of threaded discussion forums). Number each posted argument by counting the number of arguments already presented by your team. Sub-arguments need not be numbered. ARG = "argument".
EXPL	Identifies a reply/message that provides additional support, explanation, clarification, elaboration of an argument or challenge.
BUT	Identifies a reply/message that questions or challenges the merits, logic, relevancy, validity, accuracy or plausibility of a presented argument (ARG) or challenge (BUT).
EVD	Identifies a reply/message that provides proof or evidence to establish the validity of an argument or challenge.

Students also identified each message by team membership by adding an “-” for opposing or a “+” for supporting team to the message labels (e.g., +ARG, -ARG). These tags enabled students to easily locate exchanges between members from opposing teams (e.g., +ARG → -BUT) and respond to the exchanges to advance their team’s position. An example is illustrated in Figure 2.

The purpose of assigning messages to specific functions was to make the links between messages explicit, thus enabling students to visualize the structure of their arguments (Jeong and Juong, *in press*). The labels also enabled the investigator to establish each message as a unit of analysis so that message-response sequences could be clearly identified to determine their relative frequencies. Previous studies in CMC were unable to successfully measure message-response sequences (Gunawardena, Lowe, and Anderson, 1997; Newman, Johnson, Cochrane, and Webb, 1996; Levin, Kim, and Riel, 1990; Rourke, Anderson, Garrison, and Archer, 2001) because messages often addressed multiple functions at the same time. As a result, mapping the relationships between messages and replies was a difficult, if not impossible, task. In this study, message labeling was found to be an effective solution to resolving some of the problems in establishing the unit of analysis. Although these procedures may appear to be artificial and perhaps intrusive, this method has been implemented in a number of computer-supported collaborative argumentation (CSCA) systems to facilitate argumentation and problem solving (Carr and Anderson, 2001; Cho and Jonassen, 2002; Duffy, Dueber, and Hawley, 1998; Jonassen and Remidez, 2002; McAlister, 2003; Sloffer, Dueber, and Duffy, 1999; Veerman, Andriessen, and Kanselaar, 1999).

Figure 2. Example of online debate with labeled messages

<input type="checkbox"/>	<input type="checkbox"/> <u>SUPPORT statement because...</u>	Student names	Sat Oct 2, 2004 11:18 am
<input type="checkbox"/>	<input type="checkbox"/> <u>+ARG#1 MedialsButAMereVehicle</u>	Student names	Mon Oct 4, 2004 8:47 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>-EVID MedialsButAMereVeh...</u>	Student names	Tue Oct 5, 2004 7:09 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>+But RelativityTheory...</u>	Student names	Tue Oct 5, 2004 9:43 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>-But RelativityThe...</u>	Student names	Sat Oct 9, 2004 10:12 am
<input type="checkbox"/>	<u>-BUT Whataboutemotions?</u>	Student names	Tue Oct 5, 2004 9:53 pm
<input type="checkbox"/>	<u>+EVID DistEdEffectiveAsF2F</u>	Student names	Tue Oct 5, 2004 10:40 pm
<input type="checkbox"/>	<u>-BUTMediaamerevehicle</u>	Student names	Wed Oct 6, 2004 8:19 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>+EVID MooreConcurs</u>	Student names	Wed Oct 6, 2004 10:07 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>+EXPLMediaSelectionCo...</u>	Student names	Sun Oct 10, 2004 12:35 am
<input type="checkbox"/>	<input type="checkbox"/> <u>-BUT WellChosenEffect...</u>	Student names	Sun Oct 10, 2004 4:31 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>+But SupportingRes...</u>	Student names	Sun Oct 10, 2004 5:37 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>-BUTMediaismorethanamere...</u>	Student names	Fri Oct 8, 2004 5:30 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>+BUT SupportingEviden...</u>	Student names	Sat Oct 9, 2004 8:51 am
<input type="checkbox"/>	<input type="checkbox"/> <u>-BUT LearningNotSimplyAP...</u>	Student names	Mon Oct 11, 2004 9:54 am
<input type="checkbox"/>	<input type="checkbox"/> <u>+ARG2 Standards for teaching</u>	Student names	Wed Oct 6, 2004 1:48 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>+But Clarification?</u>	Student names	Sun Oct 10, 2004 5:39 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>+ARG3 MediaUnrelatedtoLearn...</u>	Student names	Wed Oct 6, 2004 3:12 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>-BUTMediaUnrelatedtoLearn...</u>	Student names	Wed Oct 6, 2004 8:26 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>+BUT MediaSelection</u>	Student names	Thu Oct 7, 2004 9:20 am
<input type="checkbox"/>	<input type="checkbox"/> <u>-BUT MediaSelection</u>	Student names	Sun Oct 10, 2004 11:21 am
<input type="checkbox"/>	<input type="checkbox"/> <u>+EVID MethodNotMedia</u>	Student names	Wed Oct 6, 2004 11:04 pm
<input type="checkbox"/>	<input type="checkbox"/> <u>-BUT MediaUnrelatedtoLea...</u>	Student names	Sat Oct 9, 2004 10:59 am

Data preparation

Computer software was written by the investigator to download and compile messages from *Blackboard* discussion forums into *Microsoft Excel*. The codes assigned to each message by the students were automatically pulled from the subject headings to identify each message an argument (ARG), evidence (EVID), challenge (BUT), or explanation (EXPL). The seven most commonly used qualifiers (“but”, “if”, “may/ might”, “I think”, “often”, “probably”, and “though”) identified by Fahy (2002b) were used to assign messages to the qualifiers group. The five most commonly used intensifiers (“very”, “only”, “every”, “never”, and “always”) were used to assign messages to the intensifiers group.

Table 1 shows the frequencies of each indicator observed in this study and in Fahy’s study. The presence of any of the select indicators found within the message text determined which messages were assigned to which linguistic group. As a result, messages were coded into four groups – (1) messages with qualifiers, (2) with intensifiers, (3) messages with neither, and (4) messages with both qualifiers and intensifiers. Tables 2 and 3 shows the mean number of replies elicited by arguments and challenges, respectively, presented by group and gender.

Inter-rater reliability

One debate from each course was randomly selected and coded by the investigator to test for errors in students’ message labels. Overall percent agreement was .91 based on the analysis of codes assigned to 158 messages consisting of 42 arguments, 17 supporting evidence, 81 critiques, and 17 explanations. The Cohen Kappa coefficient, which accounts for chance in coding errors based on the number of categories in the coding scheme, was .86 – indicating excellent inter-rater

reliability given that Kappa values of .40 to .60 is considered fair, .60 to .75 as good, and over .75 as excellent (Bakeman and Gottman, 1997, p. 66).

Table 1. Frequency of qualifiers and intensifiers observed in study

	Words	<i>n</i>	#Msgs	%Msgs	%	Fahy%
Qualifiers	But	308	242	31%	37%	34%
	If	206	167	21%	25%	23%
	May/might	180	145	18%	22%	12%
	I think	45	41	5%	5%	12%
	Often	20	17	2%	2%	9%
	Probably	8	8	1%	1%	5%
	Though	68	64	8%	8%	5%
		835	453	58%	74%	
Intensifiers	Very	72	67	9%	25%	37%
	Only	128	110	14%	45%	26%
	Every	43	38	5%	15%	22%
	Never	16	13	2%	6%	8%
	Always	28	26	3%	10%	8%
	Total	287	221	28%	26%	28%
	Total	1122		786		

Note: Females used 43 percent more qualifiers per message ($M = 1.21$, $STD = 1.49$, $n = 470$) than males ($M = .84$, $STD = 1.04$, $n = 312$). Females used 22 percent more intensifiers per message ($M = .39$, $STD = .68$, $n = 469$) than males ($M = .32$, $STD = .62$, $n = 312$).

Table 2. Mean number of replies elicited by arguments presented by group and by gender

Group	Gender	<i>Mean</i>	<i>STD</i>	<i>n</i>
Qualifiers	Female	.86	1.06	29
	Male	1.50	1.34	18
	Total	1.11	1.20	47
Intensifiers	Female	1.00	.63	11
	Male	2.80	1.48	5
	Total	1.56	1.26	16
Neither	Female	1.89	1.33	44
	Male	1.89	1.11	38
	Total	1.89	1.23	82
Total	Female	1.42	1.26	84
	Male	1.85	1.24	61
	Total	1.60	1.27	145

Note: Based on messages posted by 22 females and 10 males

Table 3. Mean number of replies elicited by challenges presented by group and by gender

Group	Gender	Mean	STD	n
Qualifiers	Female	.44	.62	107
	Male	.66	.70	70
	Total	.53	.66	177
Intensifiers	Female	.61	.61	18
	Male	.85	.55	13
	Total	.71	.59	31
Neither	Female	.53	.75	59
	Male	.66	.77	61
	Total	.59	.76	120
Total	Female	.48	.66	184
	Male	.67	.72	144
	Total	.57	.69	328

Note: Based on messages posted by 22 females and 10 males

Statistical Analysis

A 3 (linguistic form) x 2 (gender) univariate analysis of variance was used to test for differences in the mean number of replies elicited per *argument* (the dependent variable) across two independent variables – linguistic form (qualifiers versus intensifiers versus neither) and gender (males versus females). The same analysis was used to test for differences in the mean number of replies elicited per *challenge*. Arguments and challenges containing *both* qualifiers and intensifiers were not tested because these messages often contained more qualifiers than intensifiers or vice versa, and thus, interpreting their precise effects would be problematic. The effects of messages presented with both qualifiers and intensifiers will be addressed separately in another study.

To test for differences in the *distribution and patterns in replies to arguments* presented with qualifiers, intensifiers, and neither, a three-sample Chi-square test of independence was used. Similarly, a three-sample Chi-square test of independence was used to test for differences in the distribution of replies to *challenges* presented with qualifiers, intensifiers, and neither. The purpose of these tests were to determine which particular linguistic forms were more likely to produce discourse patterns that are most likely to generate sequences of speech acts that produce higher levels of critical analysis (e.g., argument → challenge → explain).

Given the exploratory nature of this study, the experiment-wise error was set at alpha level $p = .10$. As a result, each of the four tests (two ANOVA and two Chi-square tests) were conducted at $p = .10 / 4 = .025$. Note that the frequency of male arguments with intensifiers was only $n = 5$ (see Table 2). Nevertheless, this data was included in the first ANOVA test because: (a) the overall trend in the differences in number of replies elicited by arguments between linguistic groups was consistent with the differences observed in the number of replies elicited by challenges, supporting evidence, and explanations between linguistic groups (see reply rates in Figure 3); and

(b) the investigator chose to take a more liberal approach in order to fully explore and identify issues for future study.

Figure 3. Transitional probability matrices produced by the Discussion Analysis Tool containing the response distributions for messages by category X group

Transitional probabilities for messages using qualifiers

	ARG	BUT	EVI	EXP	Replies	No Replies	Givens	Reply Rate
ARGq	.00	.67	.15	.17	52	19	47	.60
BUTq	.00	.80	.11	.09	95	100	177	.44
EVIq	.00	.85	.08	.08	13	28	37	.30
EXPq	.00	.75	.12	.12	16	15	29	.48
	0	134	21	21	176	160	290	.45

Transitional probabilities for messages using intensifiers

	ARG	BUT	EVI	EXP	Replies	No Replies	Givens	Reply Rate
ARGi	.00	.64	.20	.16	25	2	16	.87
BUTi	.00	.62	.21	.17	24	11	31	.65
EVIi	.00	.80	.00	.20	5	4	8	.50
EXPi	.00	.75	.00	.25	4	1	4	.75
	0	38	10	10	58	18	59	.69

Transitional probabilities for messages using neither qualifiers nor intensifiers

	ARG	BUT	EVI	EXP	Replies	No Replies	Givens	Reply Rate
ARGn	.01	.54	.30	.15	155	9	82	.89
BUTn	.00	.68	.14	.18	71	67	120	.44
EVI n	.00	.62	.14	.24	21	25	42	.40
EXPn	.00	.72	.17	.11	18	16	31	.48
	2	157	62	44	265	117	275	.56

Note: ARG = argument, BUT = challenge, EVI = supporting or counter evidence, EXP = explanation. For example, the top matrix shows that 67 percent of the 52 replies to ARGq were challenges (BUT). In contrast, the third matrix shows that 54 percent of the 155 replies to ARGn were challenges.

Results

Effects on number of elicited replies

Responses to arguments. The 3 (linguistic form) x 2 (gender) univariate analysis of variance revealed significant differences in the mean number of replies elicited by arguments presented with qualifiers versus intensifiers versus neither, $F(2, 139) = 5.41, p = .005$. Table 2 shows that the mean number of replies elicited per argument was 1.10 ($SD = 1.20, n = 47$) with qualifiers, 1.56 ($SD = 1.26, n = 16$) with intensifiers, and 1.89 ($SD = 1.22, n = 82$) with neither. The mean number of replies elicited with qualifiers was 29 percent below the mean number of replies elicited with intensifiers (effect size = -0.37), and 41 percent below the mean number of replies elicited with neither (effect size = -0.64). The mean number of replies elicited with intensifiers was 17 percent lower than the mean number of replies elicited with neither (effect size = -0.26).

Significant differences were found in the mean number of replies elicited per male versus female argument, $F(1, 139) = 9.83, p = .002$. Table 2 shows that the mean number of replies elicited per female and male argument was 1.41 ($SD = 1.26, n = 84$) and 1.85 ($SD = 1.23, n = 61$), respectively. As a result, the mean number of replies elicited per female argument was 24 percent lower than the mean number of replies elicited per male argument (effect size = -0.35).

The effects of linguistic form on the mean number of replies elicited per argument was found to significantly vary and interact with the gender of the participant posting the argument, $F(2, 139) = 3.70, p = .027$. The mean number of replies elicited per female argument was .86 ($SD = 1.06, n = 29$) with qualifiers, which increased to 1.0 ($SD = .63, n = 11$) with intensifiers, and increased yet again to 1.88 ($SD = 1.33, n = 44$) with neither. In contrast, the mean number of replies elicited per male argument was 1.5 ($SD = 1.34, n = 18$) with qualifiers, which increased to 2.8 ($SD = 1.48, n = 5$) with intensifiers, but then, dropped down to 1.89 ($SD = 1.11, n = 38$) with neither. Female arguments with qualifiers elicited 42 percent fewer replies than male arguments with qualifiers (effect size = -0.53). In contrast, female arguments with neither qualifiers nor intensifiers elicited an equal number of replies as male arguments with neither qualifiers nor intensifiers (effects size = -.00). These findings suggest that the effect of using qualifiers when presenting arguments is greatest when they are presented by females than by males. A possible factor that contributed to this finding was that females used more qualifiers ($M = 1.82, SD = 1.59, n = 29$) per argument than males ($M = 1.52, SD = .70, n = 18$).

A post-hoc test for differences in the mean number of challenges elicited per argument revealed no main effects between linguistic forms, $F(2, 139) = 1.09, p = .34$, and gender, $F(1, 139) = 2.71, p = .10$. However, the interaction between linguistic form and gender was significant, $F(2, 139) = 3.28, p = .04$. Hence, the effects of linguistic form on the number of challenges elicited per argument varied by the gender of the participant posting the argument. Table 4 shows that the mean number of challenges elicited per female argument was .59 ($SD = .73, n = 29$) with qualifiers, which increased to .73 ($SD = .65, n = 11$) with intensifiers, and increased yet again to 1.14 ($SD = 1.11, n = 44$) with neither. In contrast, the mean number of challenges elicited per male argument was 1.00 ($SD = 1.14, n = 18$) with qualifiers, which increased to 1.60 ($SD = 1.14, n = 5$) with intensifiers, but then, dropped down to .87 ($SD = .81, n = 38$) with neither. Female arguments with qualifiers elicited 41 percent fewer challenges than male arguments with qualifiers (effect size = -0.43). Female arguments with neither qualifiers nor intensifiers elicited 31 percent more challenges than male arguments with neither (effect size = +0.27).

Table 4. Mean number of challenges elicited by arguments presented by group and by gender

Group	Gender	Mean	STD	n
Qualifiers	Female	.59	.73	29
	Male	1.00	1.14	18
	Total	.74	.92	47
Intensifiers	Female	.73	.65	11
	Male	1.60	1.14	5
	Total	1.00	.89	16
Neither	Female	1.14	1.11	44
	Male	.87	.81	38
	Total	1.01	.99	82
Total	Female	.89	.97	84
	Male	.97	.95	61
	Total	.92	.96	145

Note: Based on messages posted by 22 females and 10 males

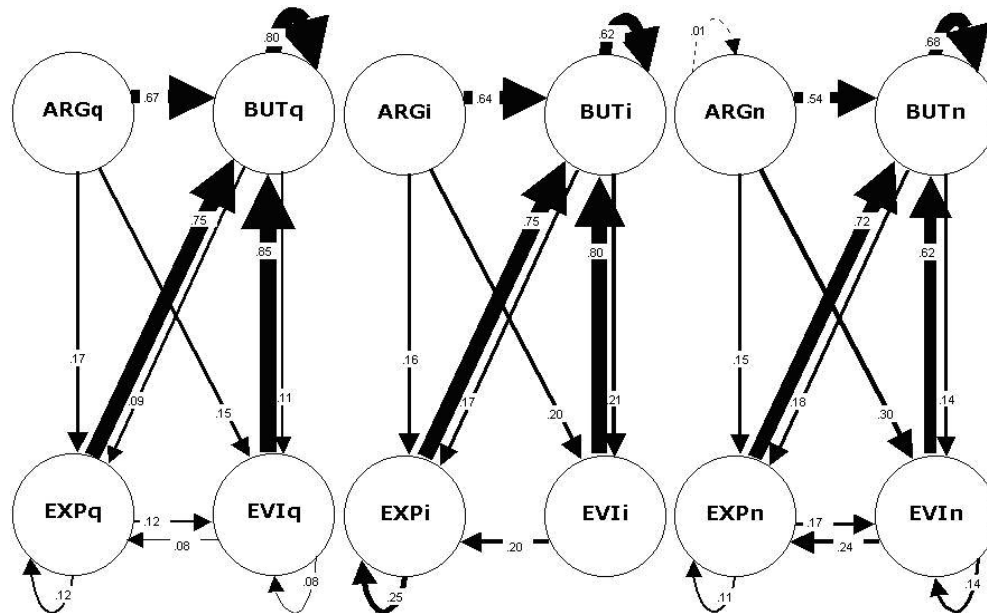
Responses to challenges

The 3 (linguistic form) x 2 (gender) univariate analysis of variance revealed no significant differences in the mean number of replies elicited per challenge with qualifiers, intensifiers, and neither, $F(2, 262) = .90, p = .41$. Table 3 shows that the mean number of replies elicited per challenge was .52 ($SD = .66, n = 177$) with qualifiers, .71 ($SD = .59, n = 31$) with intensifiers, and .59 ($SD = .76, n = 120$) with neither. The mean number of replies elicited per challenge with qualifiers was 26 percent below than the mean number of replies elicited with intensifiers (effect size = -0.29), and 11 percent below the mean number of replies elicited with neither (effect size = -0.09). The mean number of replies elicited per challenge with intensifiers was 20 percent greater than the mean number of replies elicited with neither (effect size = +0.17).

No significant differences were found in the mean number of replies elicited per male versus female challenge, $F(1, 262) = 3.79, p = .052$. The mean number of replies elicited per female and male challenge was .48 ($SD = .66, n = 184$) and .67 ($SD = .72, n = 144$), respectively. The mean number of replies elicited per female challenge was 28 percent lower than the number elicited per male challenge (effect size = -0.27).

No significant interaction was found between linguistic form and gender on the mean number of replies elicited per challenge, $F(2, 262) = .16, p = .85$. The effects of linguistic form did not depend on the gender of the participant that presented the challenges. The mean number of replies elicited per female challenge was .44 ($SD = .62, n = 107$) with qualifiers, .61 ($SD = .61, n = 18$) with intensifiers, and .52 ($SD = .75, n = 59$) with neither. In a similar pattern, the mean number of replies elicited per male challenge was .66 ($SD = .70, n = 70$) with qualifiers, which increased to .84 ($SD = .55, n = 13$) with intensifiers, but then, dropped to .66 ($SD = .77, n = 61$) with neither.

Figure 4. State diagrams of response patterns between message categories within groups



Note: ARG = argument, BUT = challenge, EVI = supporting or counter evidence, EXP = explanation. The line density reflects the transitional probabilities observed between each message pair shown in the transitional probability matrices in Figure 3. For example, the first diagram shows that 67 percent of all replies to qualified arguments were challenges, 15 percent were supporting evidence, and 17 percent were explanations.

The three matrices in Figure 3 reveal the response distributions elicited by each message category within each linguistic group. These distributions were computed using the Discussion Analysis Tool (Jeong, 2005), which tallied the frequencies and computed the relative frequencies for each observed message-response pair. The three-sample Chi-square tests revealed that differences in the response distribution elicited by: (a) ARGq versus ARGi versus ARGn were not statistically significant, $\chi(6) = 5.99, p = .42$; and (b) BUTq versus BUTi, versus BUTn were also not statistically significant, $\chi(4) = 5.4, p = .25$. The diagrams in Figure 4 are “state diagrams” (Bakeman and Gottman, 1997, p. 97) or “events networks” (Rothwell and Kazanas, 1998, p. 137) depicting the response distributions or “activity paths” triggered by each message category presented with qualifiers, intensifiers, and neither.

The diagrams suggest that qualifiers are more likely to produce (ARG → BUT) and (BUT → BUT) exchanges, and hence, more likely to produce (ARG → BUT → BUT) than messages without qualifiers. In contrast the diagrams also suggest the possibility that messages without qualifiers were more likely to produce BUT → EXP. Therefore, messages without qualifiers may be more likely to produce sequences like ARG → BUT → EXP that lead to deeper reflection and analysis of arguments. The data in this study, however, revealed that these response distributions or activity paths overall were not significantly different. As a result, one linguistic form did not necessarily produce response patterns that promoted more critical analysis than any other linguistic form.

Discussion

The purpose of this study was to explore the effects of linguistic qualifiers and intensifiers on how participants interacted and exchanged messages in online debates when using asynchronous threaded discussion forums. This study found that: (a) qualified arguments elicited significantly fewer replies (with moderate effect sizes) than arguments presented with intensifiers and neither; (b) the reduction in number of replies elicited by qualified arguments was greater when arguments were presented by females than males; and (c) the number of replies elicited by challenges with and without qualifiers was not significantly different. At the same time, however, this study also found no differences in the response patterns elicited by arguments and challenges between linguistic groups. Thus, no clear evidence was found to indicate that any one linguistic form, when used to present arguments and challenges, were more likely to trigger sequences of exchanges that promote more critical analysis of arguments than another linguistic form.

Fewer replies elicited by arguments

This study found that qualified arguments elicited fewer replies than arguments without qualifiers. One likely explanation for this finding is that the participants in this study used qualifiers primarily to hedge their claims in order to make their claims less vulnerable to criticism (given the competitive nature of the debates) in contrast to making their claims open for discussion. If qualifiers were used primarily to avoid conflict, then these findings are consistent with the Dialogic theory and its assumption that conflict (or the absence of conflict) is what drives (or inhibits) further inquiry and dialog (Bakhtin, 1981; Koschmann, 1999). Future studies, however, are needed to determine when qualifiers are used to hedge claims and when they are used to leave claims open for discussion, and how the effects of qualifiers differ when claims are presented in open exploratory discussions versus argumentative discussions.

Number of replies elicited by arguments can vary by gender

Female arguments elicited fewer replies than male arguments when presented with qualifiers. One possible explanation is that these differences may simply have been a product of the tendencies of females to engage in more supportive interactions combined with the tendencies of males to engage in more argumentative interactions with other participants (Fahy, 2002a). This could have been exacerbated by the way females in this study used 43 percent more qualifiers per message than males, thus making female arguments less likely than male arguments to elicit critical responses.

With intensifiers, the number of replies elicited by female arguments was also less than the number elicited by male arguments. One possible explanation for this finding is that the females might have been more cautious than males in using intensifiers and used them only when they felt that their argument had strong merits and hence were more difficult to refute. In contrast, the males might have used intensifiers more liberally and were more emphatic than the females when presenting both strong and weak arguments. As a result, the male arguments presented with intensifiers may have incited more criticism and thus elicited more replies.

In contrast, female arguments elicited a nearly equal number of replies as male arguments when they were stated factually (with neither qualifiers nor intensifiers). One possible explanation for this finding is that some females may have been perceived to be less credible than males simply based on their gender (Bradley, 1981). Thus, the absence of qualifiers or hedges in the female arguments made them more susceptible to challenges and questioning than male arguments

without qualifiers. All these explanations are purely speculative at this time, and will require further investigation.

No differences in number of replies elicited by challenges

In this study, no differences were found in the number of replies elicited by challenges presented with versus without qualifiers. One possible explanation for this finding is that qualifiers may have been used primarily to qualify an opposing argument (or to point out the conditions that limit the merits and plausibility of an argument), not to qualify the ideas or content presented in the challenge in itself. The other possible explanation is that the act of posting a challenge in reply to an argument initiated the conflict needed to drive further inquiry (based on the assumptions of the Dialogic theory), and that resulting drive compensated for the inhibiting effects of qualifiers (or decreases in number of replies elicited by qualified arguments).

Effects of linguistic form on response patterns and sequences

This study did not find clear evidence to indicate that any one linguistic form, when used to present arguments and challenges, were more likely to trigger response patterns and sequences of exchanges that lead to higher levels of critical analysis than another linguistic form. Nevertheless, the state diagrams generated in this study serve as useful tools for identifying and predicting potential areas where linguistic qualifiers and intensifiers could make a potential impact (when used in other contexts) on how participants interact and engage in critical discourse. One response pattern to examine more closely in future investigations is how the challenges stated with qualifiers appeared in this study to trigger a disproportionately high number of counter-challenges. What effect this level of contention has on eliciting subsequent replies, and how subsequent replies contribute to increased depth in critical analysis and discussion will require further investigation.

Instructional Implications

Overall, the findings in this study, although not conclusive, suggest that qualifiers were used primarily to hedge arguments in the online debates, and as a result, qualifiers tended to decrease the number of replies elicited by arguments. This finding suggests that the use of qualifiers should be discouraged during the initial stages of identifying arguments in order to: (a) avoid precluding others from sharing opposing viewpoints to challenge arguments; (b) elicit more diverse viewpoints and reactions; (c) support a more thorough analysis of arguments; and (d) maximize opportunities to achieve new insights and understanding. To some extent, this is similar to the rule often applied in group brainstorming where participants are asked to refrain from critiquing proposed arguments (including one's own claims) until all arguments are presented. The findings in this study suggest that implementing such a rule can lead to a substantial increase (with moderate effect sizes) in replies to arguments, and increases in level of critical analysis. The greatest gains might be achieved when this rule is applied in mostly female or all-female groups given the interaction between linguistic form and gender observed in this study.

Directions for Future Research

Once again, the findings in this study are not conclusive due to the exploratory nature of this investigation and due to particular situational variables surrounding the online debates. To conduct a closer examination of the findings reported in this study and to address some of the limitations of this study, future studies will need to: (a) examine a broader range of linguistic

phrases to better discriminate messages between linguistic groups; (b) analyze a larger sample of messages that use intensifiers; (c) identify the combination of words that can be used to discriminate when qualifiers are used to hedge claims from when they are used to leave claims open for discussion; (d) test the effects of qualifiers in the context of other group tasks and goals; (e) observe a larger number of discussion groups to prevent any idiosyncrasies in the social dynamics of any one group from potentially skewing the findings; (f) observe groups with different gender compositions and group size; and (g) examine the effects of linguistic qualifiers and intensifiers used in different contexts with and without message constraints.

In conclusion, this study provides a preliminary glimpse into the combined effects of message function, linguistic form, and gender on group interaction and group performance in CMC, and how particular patterns of interaction support critical discussions. The methods and tools described in this study will hopefully serve as a framework for investigating the effects of other communication styles and linguistic forms such as emoticons, humor, and rhetorical questions that can potentially affect both the form and function of messages and their ability to promote critical analysis, reflection, and the construction of shared meanings. Measuring the combined effects of message function and form will hopefully enable future researchers and instructional designers to develop more precise strategies for sequencing speech acts to optimize group performance in collaborative work, problem-solving, and learning in computer-mediated environments.

References

- Austin, R. (1997). Computer Conferencing: Discourse, education and conflict mediation. *Computers & Education*, 29(4), 153 – 161.
- Bakeman R., and Gottman, J. M. (1997). *Observing Interaction: An introduction to sequential analysis*. (2nd edition). University Press, Cambridge.
- Baker, M. (1999). Argumentation and constructive interaction. In P. Courier and J. E. B. Andriessen (Eds.) *Foundations of Argumentative Text Processing* (pp. 179-202). Amsterdam: Amsterdam University Press.
- Bakhtin, M. (1981). Discourse in the novel (M. Holquist & C. Emerson, Trans.) In M. Holquist (Ed.) *The Dialogic Imagination* (pp. 259-422). Austin, TX.: The University of Texas Press.
- Beers, P. J., Boshuizen, E., and Kirschner, P. (2004). Computer support for knowledge construction in collaborative learning environments. Paper presented at the *American Educational Research Association Conference 2004*, San Diego, CA.
- Blum, K. (1999). Gender Differences in Asynchronous Learning in Higher Education: Learning styles, participation barriers and communication patterns. *Journal of Asynchronous Learning Networks*, 3(1). 46 – 66.
- Bradley, P. H. (1981). The Fold-linguistics of Women's Speech: An empirical examination. *Communication Monographs*, 48, 73 – 90.
- Burgoon, J., Birk, T., and Pfau, M. (1990). Nonverbal behaviors, persuasion, and credibility. *Human Communication Research*, 17(1), 140 – 169.

- Carr, C., and Anderson, A. (2001). Computer-supported Collaborative Argumentation: Supporting problem-based learning in legal education. Paper presented at the *Computer Support for Collaborative Learning (CSCL) 2001 Conference*. Retrieved October 30, 2003 from: <http://www.mmi.unimaas.nl/euro-cscl/Papers/25.pdf>
- Cho, K., and Jonassen, D. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development*, 50 (3), 5 – 22. Retrieved March 3, 2004 from: <http://tiger.coe.missouri.edu/~jonassen/Argumentation.pdf>
- Duffy, T. M., Dueber, B., and Hawley, C. L. (1998). Critical Thinking in a Distributed Environment: A pedagogical base for the design of conferencing systems, In C. J. Bonk, and K. S. King (Eds.) *Electronic Collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse* (51-78). Mahwah, NJ.: Erlbaum.
- Fahy, P. (2002a). Epistolary and Expository Interaction Patterns in a Computer Conference Transcript. *Journal of Distance Education*, 17(1), 20 – 35.
- Fahy, P. (2002b). Use of Linguistic Qualifiers and Intensifiers in Computer Conference. *The American Journal of Distance Education*, 16(1), 5 – 22.
- Fahy, P. (2003). Indicators of Support in Online Interaction. *International Review of Research in Open and Distance Learning*, 4(1). Retrieved April 21, 2004 from: <http://www.irrodl.org/content/v4.1/fahy.html>
- Gunawardena, C., Lowe, C., and Anderson, T. (1997). Analysis of Global Online Debate and the Development of an Interaction Analysis Model for Examining Social Construction of Knowledge in Computer Conferencing. *Journal of Educational Computing Research*, 17(4), 397 – 431.
- Herring, S. (1993). Gender and democracy in computer-mediated communication. *Electronic Journal of Communication*, 3(2). Retrieved August 22, 2001 from: <http://www.cios.org/www/ejc/v3n293.htm>
- Herring, S. (1996). Two variants of an electronic message schema. In S. Herring (Ed.) *Computer-Mediated Communication: Linguistic, social and cross-cultural perspectives*, (81-106). Amsterdam: John Benjamins.
- Hosman, L. (1989). The Evaluative Consequences of Hedges, Hesitations, and Intensifiers: Powerful and powerless speech styles. *Human Communication Research*, 15(3), 383 – 406.
- Jeong, A. (2003). The sequential analysis of group interaction and critical thinking in on-line threaded discussions. *The American Journal of Distance Education*, 17(1), 25 – 43.
- Jeong, A. (2004a). Methods and tools for the computational analysis of group interaction and argumentation in asynchronous online group discussions. Paper presented at the *2005 Learning Technology Symposium*, New York, NY. Retrieved August 1, 2005, from: <http://garnet.fsu.edu/~ajeong>

- Jeong, A. (2004b). The combined effects of response time and message content on growth patterns of discussion threads in computer-supported collaborative argumentation. *Journal of Distance Education*, 19(1), 36 – 53. Retrieved August 1, 2005 from: http://cade.athabasca.ca/vol19.1/JEONG_article.pdf
- Jeong, A. (2005). *Discussion Analysis Tool*. Retrieved May 10, 2005 from: <http://garnet.fsu.edu/~ajeong/DAT>
- Jeong, A., and Juong, S. (*in press*). The effects of response constraints and message labels on group interaction and argumentation in online discussions. *Computers & Education (in press)*.
- Johnson, D., and Johnson, R. (1992). *Creative Controversy: Intellectual challenge in the classroom*. Edina, MN.: Interaction Book Company.
- Jonassen, D., and Remidez, H. (2002). Mapping alternative discourse structures onto computer conference. Paper presented at the Paper presented at *Computer Support for Collaborative Learning 2002 Conference: Foundations for a CSCL Community*, Boulder, CO.
- Koschmann, T. (1999). Paradigm shifts and instructional technology: An introduction. In T. Koschmann (Ed.) *CSCL: Theory and practice of an emerging paradigm* (1-24). Mahwah, NJ.: Lawrence Erlbaum Associates.
- Leinonen, T., Virtanen, O., and Hakkarainen, K. (2002). Collaborative discovering of key ideas in knowledge building. In Proceedings of the *Computer Support for Collaborative Learning 2002 Conference*. Boulder, CO. Retrieved May 19, 2004 from: <http://fle3.uiah.fi>
- Lemus, D., Seibold, D., Flanagan, A., and Metzger, M. (2004). Argument and decision making in computer-mediated groups. *Journal of Communication*, 54(2), 302 – 320.
- Levin, J., Kim, H., and Riel, M. (1990). Analyzing instructional interactions on electronic message networks. In L. Harasim (Ed.) *Online Education*, (pp. 185-213), New York, NY.: Praeger.
- McAlister, S. (2003). *Assessing Good Argumentation*. Retrieved April, 10, 2004, from: <http://iet.open.ac.uk/pp/s.r.mcalister/personal/AssessingGEA.htm>
- Mehrabian, A. (1968). Communication without words. *Psychology Today* 2(9), 52 – 55.
- Newman, D., Johnson, C., Cochrane, C., and Webb, B. (1996). An Experiment in Group Learning Technology: Evaluating critical thinking in face-to-face and computer supported seminars. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century*, 4(1), 57 – 74.
- Rothwell, W., and Kazanas, H. C. (1998). *Mastering the instructional design process* (2nd edition). San Francisco: Jossey-Bass.
- Rourke, L., Anderson, T., Garrison, D. R., and Archer, W. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education*, 12, 8 – 22.

- Savicki, V., Kelley, M., and Ammon, B. (2002). Effects of training on computer-mediated communication in single or mixed gender small task groups. *Computers in Human Behavior*, 18(3), 257 – 269.
- Sloffer, S., Dueber, B., and Duffy, T. (1999). *Using Asynchronous Conferencing to Promote Critical Thinking: Two implementations in higher education*. Retrieved October 30, 2003 from: <http://crlt.indiana.edu/publications/crlt99-8.pdf>
- Stegmann, K., Weinberger, A., Fischer, F., and Mandl, H. (2004). Can computer-supported cooperation scripts facilitate argumentative knowledge construction? Paper presented at the *American Educational Research Association Conference*, San Diego, CA. Retrieved May 24, 2004 from: <http://home.emp.paed.uni-muenchen.de/~weinberg/download>
- Toulmin, S.E. (1958). *The Uses of Argument*. Cambridge: University Press.
- Veerman, A., Andriessen, J., and Kanselaar, G. (1999). Collaborative learning through computer-mediated argumentation. In C. M. Hoadley and J. Roschelle (Eds.), *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference* (640-650). Palo Alto, CA.: Stanford University. Available from Lawrence Erlbaum Associates, Mahwah, NJ.
- Walther, J. (1992). Interpersonal Effects in Computer-mediated Interaction. *Communication Research*, 19(1), 52 – 90.
- Weinberger, A., Ertl, B., Fischer, F., and Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science*, 33(1), 1 – 30.
- Wiley, J., and Voss, J. (1999). Constructing Arguments from Multiple Sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, 91(2), 301 – 311.

