CrossMark

A sequential analysis of responses in online debates to postings of students exhibiting high versus low grammar and spelling errors

Allan Jeong[1] · Haiying Li[2] · Andy Jiaren Pan[3]

Abstract Given that grammatical and spelling errors have been found to influence perceived competence and credibility in written communication, this study examined how a student's grammar and spelling errors affect how other students respond to the student's postings in four online debates hosted in asynchronous threaded discussions. Message-response exchanges were sequentially analyzed to identify patterns in students' replies to arguments and challenges with counter-challenges, explanations, and evidentiary support posted by students that exhibited low versus high number of grammatical and spelling errors. Although no causal inferences can be drawn from this study, the findings nevertheless suggests that: (a) arguments posted by high-error students are more likely to be challenged than arguments posted by low-error students; (b) exchanges between high-error students can amplify the effects of grammar/spelling errors; and (c) higher levels of argumentation can be achieved by placing students into groups that are heterogeneous in writing skills in general. The findings and methods used in this study lay the groundwork for further research on strategies for managing individual differences in students' grammar and spelling (and other student behaviors in general) and increasing the level of critical discourse in online discussions.

✉ Dr. Allan Jeong
ajeong@fsu.edu

Haiying Li
haiyinglit@gmail.com

Andy Jiaren Pan
andy.c.pan3@gmail.com

[1] Instructional Systems Program, Florida State University, Tallahassee, FL, USA

[2] University of Memphis, Memphis, TN, USA

[3] Florida State University, Tallahassee, FL, USA

Springer

Introduction

Asynchronous online discussions are widely used as a key component of online courses (Ertmer et al. 2007) based on the assumption that dialogue is a process that helps students articulate supporting and opposing viewpoints, explain and justify their reasoning, and construct new meaning and knowledge (Pena-Shaff and Nicholls, 2004). However, studies on student interactions in online discussions more often than not report interactions that exhibit very little critical discourse and higher-order thinking (Meyer 2003; Rourke and Kanuka 2007; Hew et al. 2010). Moreover, most people in general are not very good arguers (Tannen 1998). One approach to addressing this problem is to use asynchronous discussion forums in ways that facilitate collaborative argumentation—an instructional activity used to foster critical reflection and discussion in both face-to-face and online environments (Veerman et al. 2002). 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, and resolve differences of opinion (Van Eemeren et al. 2008). This process not only plays a key role in increasing students' understanding but also in improving group decision-making and problem solving (Lemus et al. 2004; Wegerif et al. 1999). As a result, argumentation is a skill that is not only essential across academic disciplines like science, law, business, but also essential for everyday life.

A variety of online environments have been developed to facilitate argumentation using argumentation scripts presented in the form of question prompts and sentence openers (e.g., "I disagree because…" and "This is true because…"). This technique is used to provide students with guidance on what dialog moves to perform and at what specific times to help students construct and evaluate one another's arguments (Noroozi et al. 2012). Some of these online environments include Belvedere (Cho and Jonassen 2002; Jonassen and Kwon 2001), CSILE (Scardamalia and Bereiter 1996), ACT (Duffy et al. 1998; Sloffer et al. 1999), Hermes (Karacapilidis and Papadias 2001), FLE3 (Leinonen et al. 2002), AcademicTalk (McAlister et al. 2004), Interloc (Ravenscroft et al. 2008), and NegotiationTooli (Beers and Boshuizen 2004). The use of sentence openers for example has been found to increase the frequency of relevant claims and arguments (Schwarz and Glassner 2007), reasoned claims, disagreements, rebuttals (McAlister et al. 2004), qualified claims, counter arguments (Stegmann et al. 2007), and on-task contributions (Baker and Lund 1997; Schwarz and Glassner 2007). Overall, this body of research provides insights into the intersection between two of three elements in the "community of inquiry" framework (Garrison and Archer 2003) used to guide research on online learning—these two elements being teaching presence (e.g., designs for facilitating discourse) and cognitive presence (e.g., cognitive operations like those used in argumentation). Social presence (e.g., socio-emotional relationships, effective communication) is the third element in the framework that has been examined frequently in recent research on online learning, but one that needs to be examined in relation to (not artificially separate from) cognitive and teacher presence (Garrison 2007).

One way to gain insight into the intersection between all three elements (social, cognitive, and teaching presence) is to examine *how* students express their arguments, how this affects the way students perceive one another, and how their perceptions of one another affect the way they respond to one another in terms of the cognitive operations they perform in producing a response. One approach to conducting this type of study is to use sequential analysis to examine argumentation in terms of dialog move sequences and to

determine how particular social characteristics of a message (or messenger) influences the way students respond to the message. By using sequential analysis to capture the dynamic nature of online discourse in terms of temporal and sequential patterns in message-response exchanges, one can test and determine the relationships between social and cognitive presence (Kapur 2011). However, only two studies have used this particular approach to examine students' interactions in computer-supported collaborative argumentation (Jeong 2005a, b, 2006).

Jeong (2006) used sequential analysis to examine how messages presented versus not presented in a more pleasant, socially-amenable, and "conversational" style (with greetings, emoticons, acknowledgments, addressing others by name) affected how students responded to messages. In this study, students participated in online debates and were instructed to restrict and tag each of their postings as either an argument that supports or opposes the main claim under debate (+ARG, -ARG), a challenge (BUT), an explanation (EXPL), or supporting evidence (EVID). The findings showed that arguments presented in a conversational style elicited challenges in .90 of its replies whereas arguments presented in a purely expository style elicited challenges in only .52 of its replies. Challenges presented in a conversational style elicited up to eight times more responses providing further explanations for a given argument than challenges presented in expository style. In addition, the probability of eliciting an explanation in response to a challenge was .09 when challenges were presented in expository style versus .23 when presented in a conversational style.

In the second study, Jeong (2005a) again used sequential analysis to examine the number and types of replies elicited by arguments and challenges presented with versus without the use of linguistic qualifiers (e.g., if, maybe, might, sometimes, but, probably, I think, though) in online debates. Like Jeong's (2006) study, students participating in the debate classified and labeled their messages as arguments, challenges, supporting evidence, or explanations prior to posting each message. Using this approach, the study found that qualified arguments elicited .41 fewer replies (effect size = −.64). Furthermore, the reduction in replies was greatest when qualified arguments were presented by females as opposed to arguments posted by males. No significant differences were found in the number or the types of replies elicited by challenges presented with versus without qualifiers. Both studies above (Jeong 2005a, b, 2006) illustrate how the manner and form in which ideas are presented can influence the level of argumentation in online debates when argumentation is examined in terms of message-response sequences.

Another way in which students' postings can vary in manner and form is the number of grammatical and spelling errors (GSE) found in students' postings. At this time, no studies have yet been conducted to provide direct evidence to show a direct link between GSE and the way students post replies to messages. However, findings from prior studies have found that GSE can affect perceived credibility (Liu 2004; Morris et al. 2012) much like the way a speaker's verbal fluency, such as the number of verbal stumbles or "speed and ease of verbal production" (Lezak et al. 2004, p. 518), can affect how others perceive the competence, credibility and persuasiveness of the speaker (Burgoon et al. 1990; Berger 1997; Myers and Bryant 2004). Most of all, perceived credibility has been found to affect message acceptance (Carpenter 2012). These two sets of findings when taken together suggest that GSE can affect message acceptance. Under the assumption that accepted messages are likely to elicit responses that provide supporting evidence and explanations whereas rejected messages are likely to elicit challenges and/or counter-evidence, it is reasonable to believe that GSE in a message can potentially affect the types of responses elicited by messages.

As a result, the purpose of this study was to examine how students' GSE in online postings influence how postings elicit certain types of responses (e.g., how likely students are to

respond to arguments with challenges and respond to challenges with counter-challenges versus explanations versus supporting evidence). This study hypothesized that: (a) arguments are more likely to elicit replies that challenge the arguments when arguments are posted by students with high GSE than by students with low GSE; (b) challenges posted by students with high GSE are more likely to elicit counter-challenges, supporting evidence, and explanations than challenges posted by students with low GSE. The Dialogic theory of language (Bakhtin 1981; Koschmann 1999) provided the rationale in this study to focus exclusively on four and *only four* specific types of messages-response sequences (in order to minimize chances of committing Type I error) and how they are affected by students' GSE. The main assumption of the Dialogic theory is that conflict is produced not by examining an utterance by itself (e.g., comparing the frequency of challenges posted in an online debate), but by examining the relationship between utterances (e.g., the probability that a presented claim elicits a response that challenges the claim). As a result, meaning is re-negotiated and re-constructed as a direct result of cognitive conflict produced in social exchanges (ARG → BUT), and that conflict is the primary force that drives the processes of critical inquiry by eliciting supporting evidence (BUT → EVID), further explanations (BUT → EXPL), and rebuttals (BUT → BUT). In other words, the need to explain, justify, and understand is felt and acted upon only when conflicts or errors are brought to attention (Baker 1999). Support for this theory and its assumptions can be found from the extensive research on collaborative learning that show conflict and the consideration of both sides of an issue is what drives inquiry, reflection, articulation of individual viewpoints and underlying assumptions, and deeper understanding (Johnson and Johnson 2009; Wiley and Voss 1999).

Based on the assumptions of the Dialogic Theory and the role of conflict in driving further inquiry, the purpose of this study was to examine the effects of students' GSE on four specific message-response exchanges (ARG → BUT, BUT → BUT, BUT → EXPL, BUT → EVID) to determine how likely a presented argument elicited responses that challenged the argument, and how likely a challenge elicited supporting evidence versus explanations versus counter challenges. As a result, the questions examined in this study were the following:

1. What are the response probabilities across the four types of exchanges when arguments and challenges are posted by students that exhibit high versus low GSE, and what differences exist in the response probabilities between the two groups?
2. To explore the potential effects of GSE in *homogeneous* groups, what are the differences in response probabilities in each of the four types of exchanges when *both* the message and the responses are posted by students with high versus low GSE?
3. To explore the potential differences between homogeneous versus *heterogeneous* groups, what are the response probabilities in each of the four types of exchanges when high-GSE students respond to the messages of low-GSE students, and vice versa? How do these probabilities differ from those produced in homogeneous groups?

Method

Participants

The participants were 72 graduate students (42 females, 30 males; five international students) aggregated across the fall 2005 ($n = 20$), fall 2006 ($n = 17$), spring 2007 ($n = 20$), and fall 2007 ($n = 15$) semesters of an online graduate-level course titled "Introduction to

distance learning" at a large southeastern university. The majority of the students were Master's level students majoring in instructional systems. Eight of the students were doctoral students in instructional systems.

Debate procedures

The students participated in four weekly online team debates using asynchronous threaded discussion forums in Blackboard[TM], a web-based course management system, in weeks 4, 5, 7 and 10 of the semester. The debates were structured so that: (a) student participation in the four debates contributed to .036 of the course grade; (b) for each debate, students were required to post at least four messages; (c) prior to each debate, students were randomly assigned to one of two teams (balanced by gender with half the class assigned to one team and half the class assigned to the other team) to either support or oppose a given position; and (d) students were required to vote on the team that presented the strongest arguments following each debate. The purpose of each debate was to critically examine design issues, concepts and principles in distance learning covered during the week of the debate. Students debated on the following claims: "Given the data and needs assessment, the fictitious country of NED should not develop a distance learning system", "The Dick & Carey ISD model is an effective model for designing the instructional materials for this course", "Type of media does not make any significant contribution to student learning", and "Print is the preferred medium for delivering a course study guide".

Prior to each debate, students were presented a list of four message categories (Table 1) that determined the types of postings students can contribute to the online debates. Based on Toulmin's (1958) model of argumentation, each message category was presented with a description of its function, an example posting, and an assigned message label. In each debate, students were instructed to restrict each posting to serve only one function at a time, and to insert the corresponding message label into the subject headings of each message (Fig. 1). The course instructor occasionally checked and asked students to correct errors or omissions in the message labels and subject headings. No participation points were awarded for a given debate if a student did not follow the protocol. Overall, the percentage of labels that students correctly assigned to their postings was 86.7% (based on a review of a sample of the 1557 postings contributed to the debates).

Students were also instructed to identify each message by team membership by prefacing each message label with a "-" for opposing or "+" for supporting team (e.g., +ARG, −ARG). These tags enabled students to locate the exchanges between the opposing and supporting teams during the debates (e.g., +ARG → −BUT) and respond to the exchanges to advance their team's position. An example debate forum is illustrated in Fig. 1.

The data set

The 1557 messages posted to the debates were copied and pasted into Microsoft Word[TM] and the number of grammatical errors (identified in MS Word by green underlined text) and spelling errors (identified in MS Word by red underlined text) observed in each posting was counted. Spelling errors noted in proper names, word abbreviations and message labels posted in the subject headings were not counted as an error. Although it is not known to what extent students used grammar and spelling checking outside of the Blackboard[TM] discussion forums prior to making each posting, each student was classified into either the high or low GSE group based solely on the total number of GSE each individual student

Table 1 Example instructions on how to label messages posted to the online debates

Label	Definition of label
+	If you are on the SUPPORTING team, ALL your posted messages must include the + tag before each message label
−	If you are on the OPPOSING team, ALL your posted messages must include the—tag before each message label
ARG#	ARGUMENT: Identifies a message that presents one and only one argument or reason to support your team's position. Number each posted argument by counting the number of arguments already presented by your team. Example argument supporting use of threaded discussions over use of chat rooms: +ARG2 Produces Deeper Discussions
EXPL	EXPLANATION: Identifies a response that provides additional support or sub-arguments, explanation, clarification, or elaboration: +EXPL Can Participate in Multiple Threads
BUT	CHALLENGE: Identifies a response that questions/challenges the merits, logic, relevancy, validity, accuracy or plausibility of a claim or challenge: BUT Multiple Threads Produces Cognitive Overload
EVID	EVIDENCE: Identifies a response that provides numbers, data, proof or evidence to verify or establish the validity of an argument or challenge: +EVID Discussion Threads Are 50% Longer On Average

produced in all their postings across all four debates ($M = 4.96$, $SD = 6.99$) divided by the total number of words ($M = 513.39$, $SD = 565.8$) contributed by the student. The median score was used to classify each student as high-GSE ($n = 35$, 1 international student) and low-GSE ($n = 37$, 4 international students). Overall, 53% of the errors identified by MS Word were spelling errors and 47% were grammatical errors (e.g., use of passive voice, over use of contractions, lack of precision, missing comma, missing space after period or parentheses, over use of jargon, wordiness). To examine the validity of the grammar and spelling errors identified by MS Word, the errors found in the first two debates from the fall 2005 semester (consisting of a total of 226 student postings) were cross examined. Of the 35 spelling errors identified by MS Word, all 35 were correctly identified misspellings. Of the 55 grammatical errors identified by MS Word, 49 of the 55 (or 89%) were found to be legitimate errors in grammar/style. The 6 grammatical errors that were found to be borderline errors were cases where students used phrases commonly found in conversations but not typically found in written expository text (e.g., "Point well taken", using don't versus do not).

The message labels that students assigned to each posting was modified to identify both the type of posting (ARG, BUT, EXPL, EVID) and group (i.e., H = low-GSE, L = high-GSE). An L and H was added to the end of the label to identify a message posted students that produced a low and high number of GSE, respectively. As a result, a total of eight message types were included in the data analysis (ARGL, BUTL, EXPLL, EVIDL, ARGH, BUTH, EXPLH, EVIDH). The messages that students did not explicitly or clearly label at the time of posting were assigned the code OTHER. The message labels were then entered in one single column of a spreadsheet in the order in which they appeared on the discussion forum like the one shown in Fig. 1. To the immediate right of each message label listed in the spreadsheet, the number of indentations (or thread level) preceding each message label and heading presented in the forum was entered into the adjacent column. This sequential data was copied and pasted into the Discussion Analysis Tool or DAT (Jeong 2013) to tally

⊟ SUPPORTING ARGUMENTS	~~Allen Jeong~~	12/9/07 2:27 AM
⊟ +ARG1 MediaDevelopWorldKnowledge	~~Michael Peterson~~	12/9/07 2:53 AM
⊟ +EVID1 MediaDevelopWorldKnowledge	~~Jason Free~~	12/5/07 3:58 PM
+EXPL MediaChangeLearningEnvironments	~~Michael Peterson~~	12/5/07 3:58 PM
⊟ Re: -BUT ExpNotOutcome	~~Richard MacDonald~~	12/5/07 3:58 PM
+EXPL ExpNotOutcome	~~Michael Peterson~~	12/5/07 3:57 PM
⊟ - BUTMediaDevelopWorldKnowledge	~~Lillian Eldango~~	12/9/07 2:30 AM
⊟ +BUT WeMustMakeMediaDecisions	~~Michael Peterson~~	12/5/07 3:58 PM
⊟ -BUT WeMustMakeMediaDecisions	~~Lillian Eldango~~	12/5/07 3:58 PM
+EXPL WeMustMakeMediaDecisions	~~Michael Peterson~~	12/5/07 3:58 PM
-BUT necessaryTechnology	~~April Freeman~~	12/5/07 3:58 PM
⊟ +SUPP ComputerSimulationsDevelopWorldKnowledge	~~Monica Jackson~~	12/9/07 2:30 AM
+EXPL	~~Beatriz Kush~~	12/21/07 2:38 PM
⊟ +ARG2 LearningAspectsInfluencedByMediaCharacteristics	~~James Britton~~	12/9/07 3:01 AM
⊟ -BUT LearningAspectsInfluencedByMediaCharacteristics	~~Lillian Eldango~~	12/9/07 2:34 AM
⊟ +EVID LearningAspectsInfluencedByMediaCharacteristics	~~Elena Novak~~	12/9/07 2:34 AM
+EXPL	~~Beatriz Kush~~	12/21/07 2:38 PM
-BUT LearningAspectsInfluencedByMediaCharacteristics	~~Daniel Chavas~~	12/5/07 3:58 PM
-BUT2 LearningAspectsInfluencedByMediaCharacteristics	~~Nathan Hawthorne~~	12/9/07 2:34 AM
+EVID for LearningAspects Are InfluencedByMedia	~~Monica Jackson~~	12/9/07 2:33 AM

Fig. 1 Debate with labeled messages in a Blackboard[TM] threaded discussion forum. Messages assigned with incorrect message labels (e.g., +SUPP in line 12) were omitted from analysis

the frequency in which each type of message was posted in *direct* reply to the other message types across all 1557 postings. DAT was then used to compute transitional probabilities based on the observed frequencies, compute *z*-scores to determine which probabilities are significantly higher and lower than the expected probability, and generate transitional state diagrams to visually convey observed patterns in the transitions between the four types of messages examined in this study.

Figure 2 shows that in this study, for example, a total of 112 arguments (ARGH) were posted by high-GSE students. Among the replies to ARGH, .39 of the responses were challenges from high-GSE students (.39 = 71 BUTH replies divided by 183 total replies to ARGH). In contrast, only .18 of all responses to ARGH were challenges (BUTL) from low-GSE students (.18 = 33 divided by 182 total replies to ARGH). Using *z*-scores to operationally define which message-responses sequences are to be identified as a "response pattern", the .39 probability (presented in bold font) of a BUTH response to ARGH was significantly higher than the expected probability because its *z*-score value of +4.55 was greater than the critical *z*-score of +1.96 at $p < .05$. The .18 probability (presented in bold and underlined) of a BUTL response to ARGH was significantly lower than the expected probability because its *z*-score value of −4.06 was lower than the critical *z*-score of −1.96 at $p < .05$.

	ARGH	BUTH	EVIDH	EXPLH	ARGL	BUTL	EVIDL	EXPLL	OTHER	Replies	No Replies	Givens	Reply Rate
ARGH	.02	.39	.11	.20	.01	.18	.03	.03	.03	183	22	112	.80
	.42	4.55	3.99	6.23	1.44	-4.06	-3.86	-4.77	.45				
BUTH	.00	.50	.06	.10	.00	.17	.05	.07	.04	202	172	322	.47
	-2.10	8.82	.76	.84	-.61	-4.54	-2.99	-3.47	2.26				
EVIDH	.00	.51	.16	.13	.00	.16	.02	.02	.00	45	38	73	.48
	-.93	4.06	3.10	1.22	-.27	-2.26	-1.91	-2.43	-1.04				
EXPLH	.01	.35	.06	.25	.00	.23	.03	.01	.06	69	54	107	.50
	-.23	1.87	.17	5.00	-.34	-1.42	-2.21	-3.23	2.02				
ARGL	.02	.15	.04	.03	.00	.25	.26	.23	.02	226	22	124	.82
	-.04	-4.07	-.67	-3.17	1.20	-2.02	8.01	3.56	-.56				
BUTL	.00	.18	.01	.05	.00	.44	.12	.20	.01	240	211	393	.46
	-2.33	-2.91	-3.13	-2.37	-.68	4.80	.37	2.25	-1.18				
EVIDL	.00	.14	.06	.01	.00	.44	.18	.13	.03	77	79	143	.45
	-1.23	-2.29	.47	-2.32	-.36	2.61	2.08	-.49	.19				
EXPLL	.00	.10	.03	.09	.00	.55	.04	.17	.01	92	115	190	.39
	-1.35	-3.55	-.92	.11	-.39	5.30	-2.11	.69	-.79				
OTHER	.10	.05	.01	.00	.00	.39	.12	.31	.01	140	25	93	.73
	7.72	-5.85	-2.18	-3.80	-.50	2.29	.46	5.81	-1.31				
	23	322	68	107	2	393	140	190	29	1274	738	1557	.32

Fig. 2 DAT screen shot of a table presenting the transitional probabilities for the type and number of replies to each message type with their corresponding z-scores. *Labels* in *left column* lists the messages by category, and *labels* in the *top row* lists the types of responses posted in reply to the messages listed in the *left column*. *ARGH* and *ARGL* argument posted by high-GSE and low-GSE students, respectively; *BUT* challenges; *EXPL* explanations; *EVID* supporting evidence; *Replies* total number of message posted in reply to the given message identified in the *left-most column*; *Given* the total number of messages of the message type identified in the *left-most column*; *No Replies* the number of messages of the given message type that elicited no replies at all. Transitional probabilities in *bold* and probabilities in *bold plus underlined* are significantly higher and significantly lower than expected, respectively, based on z-score values (presented in *gray*)

The *z*-score statistic was based on a formula developed by Bakeman and Gottman (1997) that takes into account not only the observed total number of responses following each given behavior, but also the marginal totals of each response observed across all behaviors. The stability of the *z*-scores, however, depends on two conditions: a) the observed transitional probability with a significant *z*-score must be based on a cell frequency of 4 or more; and b) the total number of tallies in each frequency matrix should be four to five times the total number of cells in the frequency matrix (Bakeman and Gottman 1997, p. 145). The data presented in Fig. 2 in addition to all subsequent partitioning of the data used to address the research questions examined in this study meet both of these conditions. Given the large number of *z*-score tests presented in the *z*-score matrices, the chances of committing a Type I error was minimized by focusing on only the four types of message-response exchanges (e.g., ARG → BUT, BUT → BUT, BUT → EVID, and BUT → EXPL) viewed to be fundamental to the process of argumentation in accordance to the Dialogic theory and its assumptions.

Results and discussion

Responses to messages posted by high- versus low-GSE students

Data analysis

To examine how students (high- and low-GSE students combined) responded to messages posted by high-GSE students, the observed response frequencies for ARGL, BUTL, EXPLL, and EVIDL presented in the top *left* quadrant of Fig. 2 were aggregated with the frequencies in the top *right* quadrant of Fig. 2 to produce the transitional probability and *z*-score table presented on the left side of Fig. 3. In addition, the observed response frequencies for ARGH, BUTH, EXPLH, and EVIDH in the *lower left* quadrant of Fig. 2 were aggregated with the frequencies in the *lower right* quadrant of Fig. 2 to produce the transitional probability and *z*-score table presented on the right side of Fig. 3.

The probabilities presented in each of the two tables in Fig. 3 were converted into two transitional state diagrams (Fig. 4). The state diagrams are graphical representations used to visually and efficiently identify similarities and differences in response patterns generated by messages posted by high- versus low-GSE students. The diagrams can also be used to determine whether there are more global patterns in longer sequences and chain of responses (e.g., EXPL → BUT → BUT). In the state diagrams, the black-colored arrows and the arrows drawn with dotted lines identify transitional probabilities found to be significantly higher and lower than expected, respectively, based on *z*-score tests with alpha $p = .05$. Gray-colored arrows identify probabilities that were neither greater nor lower than the expected probability. The first and second value in each node specifies number of times the given message type was observed and the number of responses posted in reply to the given message type, respectively. The size of the glow emanating from each node conveys the frequency in which the message type was observed relative to the frequency of other message types observed *within* the group.

Main findings on response probabilities

Arguments posted by high-GSE students elicited proportionally more challenges than arguments posted by low-GSE students (.58 versus .41). These findings support the hypotheses that arguments posted by high-GSE students (and the resulting decrease in

Responses to High-Error Students

	ARG	BUT	EVID	EXPL	Replies	No Replies	Givens	Reply Rate
ARGH	.03	.58	.15	.24	178	22	112	.80
	2.37	-2.11	.75	1.21				
BUTH	.00	.70	.12	.18	193	172	322	.47
	-2.02	2.26	-.63	-1.58				
EVIDH	.00	.67	.18	.16	45	38	73	.48
	-.79	.33	.98	-.97				
EXPLH	.02	.62	.09	.28	65	54	107	.50
	.23	-.53	-.99	1.38				
	6	310	63	102	481	286	614	.14

Responses to Low-Error Students

	ARG	BUT	EVID	EXPL	Replies	No Replies	Givens	Reply Rate
ARGL	.02	.41	.31	.26	222	22	124	.82
	3.02	-5.35	5.05	.85				
BUTL	.00	.62	.13	.24	237	211	393	.46
	-1.75	2.97	-3.45	.14				
EVIDL	.00	.60	.25	.15	75	79	143	.45
	-.83	.95	1.19	-2.05				
EXPLL	.00	.66	.08	.26	91	115	190	.39
	-.93	2.29	-3.21	.53				
	5	343	126	151	625	427	850	.13

Fig. 3 Screen shot from DAT with aggregated response frequencies and probabilities for examining patterns in responses to messages posted by high and low-GSE students
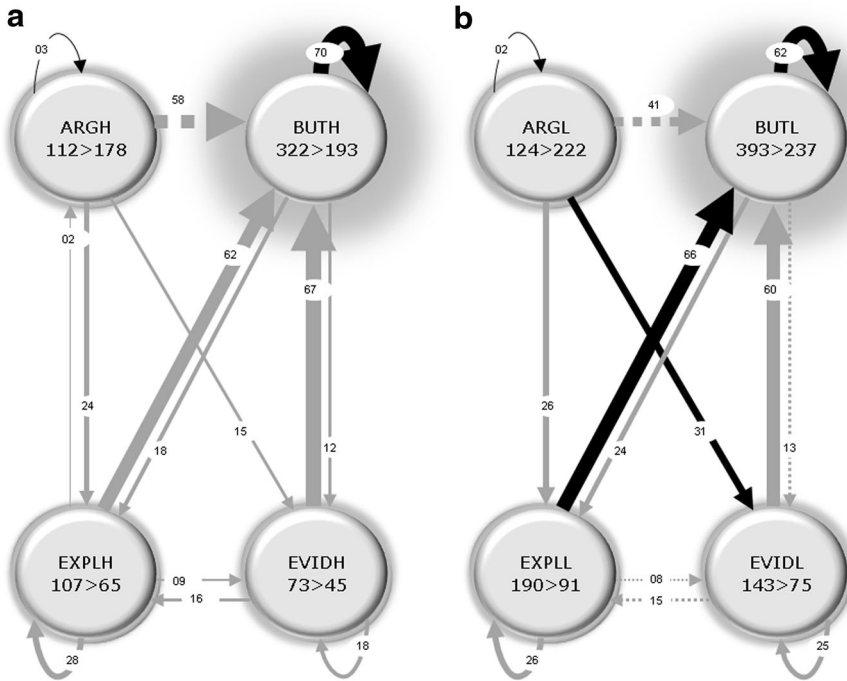
Fig. 4 Transitional state diagrams revealing patterns in responses to messages posted by high- versus low-GSE students. **a** Responses to messages posted by students with HIGH errors. **b** Responses to messages posted by students with LOW errors. *ARG* argument, *BUT* challenge, *EVID* supporting evidence, *EXPL* explanation; *H* suffix denotes messages posted by low-GSE students, *L* suffix denotes messages posted by high-GSE students; *Black* and *gray colored arrows* denote transitional probabilities that are and are not significantly higher than expected, respectively, based on *z*-scores at $p < .05$; *Dotted lines* denote probabilities significantly lower than expected; the size of glow surrounding each dialog move conveys the relative frequency of the dialog move; *first digit* in node conveys number of times dialog move was observed, and *second digit* following the *greater than symbol* is the total number of messages posted in response to the dialog move

perceived credibility/authority) can increase students' inclination to challenge the arguments posted by high-GSE students. Challenges posted by high-GSE students elicited proportionally more counter-challenges than challenges posted by low-GSE students (.70 versus .62). The findings also show that challenges posted by high-GSE students elicited proportionately fewer responses with supporting evidence than challenges posted by low-GSE students (.12 versus .13) and fewer responses to explain and build on the arguments (.18 versus .24). As a result, the overall trend in these findings provide some support for the hypotheses that challenges posted by high-GSE students (and students that are possibly perceived to possess less credibility/authority) elicits proportionately fewer rebuttals and justifications in the form of explanations and supporting evidence.

The Yule's Q test was used to determine to what extent these observed differences in response probabilities were associated with students exhibiting high and low GSE. The range of possible values for the phi-coefficient ø is −1 to 1 with zero indicating no association, with .90 an above indicating an extremely strong relationship, .70 to .89 indicating a strong relationship, .50 to 69 a moderate relationship, .30 to .49 a low relationship, and .30 and below a weak relationship. The tests revealed the following findings:

(1) differences in the probability of challenges posted in reply to the arguments of high-versus low-GSE students (.58 versus .41) showed a low association with students' errors ($\phi = .242$; $SE = .075$); (2) differences in the probability of counter-challenges posted in reply to challenges from high- versus low-GSE students (.70 versus .62) showed little or no association with student errors ($\phi = .119$; $SE = .068$); (3) differences in the probability of evidentiary support posted in reply to challenges from high- versus low-GSE students (.12 versus .13) showed no association with student errors ($\phi = -.019$; $SE = .141$); and (4) differences in the probability of explanations posted in reply to challenges from high-versus low-GSE students (.18 versus .24) showed little or no association with student errors ($\phi = -.147$; $SE = .11$).

Post-hoc tests on differences in response frequencies

Given the observed differences in some of the response probabilities, post hoc Bonferroni two-tailed t-tests were conducted to test for differences in response frequencies. A significant difference was found in the mean number of challenges posted in reply to arguments posted by high ($M = .93$, $SD = 1.03$, $n = 112$) versus low-GSE ($M = .59$, $SD = .82$, $n = 248$) students, $t(358) = -3.30$, $p = .001$. Arguments posted by high-GSE students elicited .57 more challenges ($ES = + .18$) than arguments posted by low-GSE students. No significant difference was found in the mean number of counter challenges posted in reply to challenges from high-GSE ($M = .42$, $SD = .689$, $n = 322$) versus low-GSE students ($M = .38$, $SD = .631$, $n = 393$), $t(713) = -.925$, $p = .36$. No significant difference was found in the mean number of EVID posted in reply to BUTH ($M = .071$, $SD = .258$, $n = 322$) versus BUTL ($M = .079$, $SD = .279$, $n = 393$), $t(713) = .367$, $p = .71$. No significant difference was found in the mean number of EXPL posted in reply to BUTH ($M = .10$, $SD = .318$, $n = 322$) versus BUTL ($M = .15$, $SD = .367$, $n = 393$), $t(713) = -1.59$, $p = .11$. These findings altogether correspond with the observed differences in response probabilities because: a) higher response probabilities correlated with higher response frequencies (Pearson $r = .734$); and b) significant differences in response frequencies were found where some association was found between response probabilities and errors, and no significant differences in response frequencies were found where little or no association was found between response probabilities and errors.

Overall comparisons of message-response patterns

A full post hoc comparison of the state diagrams in Fig. 6 revealed two *similar* patterns in the responses to messages posted by low-GSE and high-GSE students: (1) arguments from both high- and low-GSE students had the tendency to elicit *less* than the expected number of challenges, with arguments from high-GSE students eliciting challenges in .58 of its responses (z-score $= -2.11$, $p < .05$) and arguments from low-GSE students eliciting challenges in .41 of its responses (z-score $= -5.35$, $p < .05$); and (2) challenges from both high- and low-GSE students elicited significantly *more* than expected counter-challenges, with challenges from high-GSE students eliciting counter-challenges in .70 of its responses (z-score $= 2.26$, $p < .05$) and arguments from low-GSE students eliciting counter-challenges in .62 of its responses (z-score $= 2.97$, $p < .05$). Because both high- and low-GSE students showed a tendency to respond to arguments with challenges *less* often than expected, this likely helped to increase the potential for the arguments of high-GSE students to elicit more challenges than the arguments posted by low-GSE students (which was the case as reported above). In contrast, there was likely to be less potential for the

challenges from high-GSE students to elicit more counter challenges than the challenges from low-GSE students because the challenges from low-GSE students already possessed the tendency to elicit counter-challenges at a higher than expected probability. Note: The number of arguments posted in reply to other arguments was omitted from analysis because students were instructed to post each argument in a new discussion thread (not in a reply to a previously posted argument).

With regards to the *differences* in response patterns, the state diagrams revealed three unique patterns in the responses to messages posted by low-GSE students. Arguments posted by low-GSE students showed a tendency to elicit supporting evidence (z-score $= 5.05$, $p < .05$). Yule's Q test revealed a low association between this pattern and student errors ($\emptyset = -.369$; $SE = .103$). This finding suggests that students were more inclined to accept and lend evidentiary support (as opposed to challenging) to the arguments posted by low-GSE students (students that may be perceived to have more credibility/authority). Challenges posted by low-GSE students elicited fewer than expected responses with supporting evidence (z-score $= -3.45$, $p < .05$). This finding suggests that students were more likely to bow down to the challenges from low-GSE students. As a result, students were less inclined to rebut these challenges with evidentiary support for the challenged argument. However, the Yule's Q test revealed no association between this pattern and student errors ($\emptyset = -.019$; $SE = .141$). Explanations posted by low-GSE students had the tendency to elicit more challenges (z-score $= 2.29$, $p < .05$). However, the Yule's Q test revealed no association between this pattern and student group ($\emptyset = -.078$; $SE = .106$).

Response patterns in exchanges between students within group

Data analysis

To determine the potential effects of GSE when students engage in argumentation in homogeneous groups, patterns in how high-GSE students responded to other high-GSE students were compared with how low-GSE students responded to other low-GSE students. To make this comparison, the frequencies reported in the upper left quadrant of Fig. 2 (along with the marginal total for number of givens) were extracted and placed into its own frequency matrix in the DAT software (see Fig. 5) to compute a new set of transitional probabilities and z-scores. These results are presented graphically in the right transitional state diagram in Fig. 6. Next, the frequencies reported in the lower right quadrant of Fig. 2 (along with the marginal total for number of givens) were extracted and placed into its own frequency matrix to compute a new set of transitional probabilities and z-scores. These results are presented graphically in the left transitional state diagram in Fig. 6.

Differences in response probabilities within group

The arguments posted by high-GSE students elicited proportionally more challenges than arguments posted by low-GSE students (.53 versus .34). The Yule's Q test showed a low association between this response pattern and student errors ($\emptyset = .306$; $SE = .087$). This finding supports the hypotheses that high-GSE students (and resulting decrease in perceived credibility/authority) can increase students' inclination (particularly high-GSE students) to challenge the arguments posted by high-GSE students. Challenges posted by high-GSE students elicited proportionally more counter-challenges than challenges posted by low-GSE students (.75 versus .58). However, the Yule's Q test showed a weak

association between this response pattern and student errors (ø = .181; *SE* = .078). The challenges posted by high-GSE students elicited proportionately fewer responses with supporting evidence than challenges posted by low-GSE students (.10 versus .16). The Yule's Q test showed a weak association between this response pattern and student errors (ø = −.234; *SE* = .162). The challenges posted by high-GSE students elicited proportionately fewer explanations than the challenges posted by low-GSE students (.15 versus .26). The Yule's Q test showed a weak association between this response pattern and student errors (ø = −.285; *SE* = .127). Although the association between the patterns of responses to challenges and student errors were weak overall, the overall consistency and trends in the observed differences supports the hypothesized effects of GSE on how students respond to arguments and challenges. Furthermore, the phi coefficients observed across all four types of exchanges between students that exhibited the *same* level of errors (low or high) were consistently greater in absolute value than the phi coefficient produced in exchanges between students with mixed levels of errors (.24 vs. 30, .12 vs. .18, −.02 vs. −.23, −.15 vs. -.28). This finding suggests that the effects of GSE may be amplified when students debate in groups that are homogeneous in GSE.

Differences in response frequencies

A post hoc Bonferroni *t* test revealed no significant differences in the mean number of challenges posted in response to arguments from high-GSE students (*M* = .663, *SD* = .88, *n* = 112) versus low-GSE students (*M* = .459, *SD* = .725, *n* = 124), *t(df = 234)* = 1.665, *p* = .097. No significant differences were found in the mean number of counter-challenges posted in response to challenges from high-GSE students (*M* = .313, *SD* = .639, *n* = 322) versus low-GSE students (*M* = .266, *SD* = .536, *n* = 394), *t(df = 714)* = 1.073, *p* = .283. No significant differences were found in the mean number of evidence posted in response to challenges from high-GSE students (*M* = .04, *SD* = .197, *n* = 332) versus low-GSE students (*M* = .071, *SD* = .267, *n* = 393), *t(df = 713)* = -1.724, *p* = .085. No significant differences were found in the mean number of explanations posted in response to challenges high-GSE students (*M* = .062, *SD* = .254, *n* = 322) versus from low-GSE students (*M* = .119, *SD* = .347, *n* = 394), *t(df = 714)* = -2.464, *p* = .014. Although the correlation between these response frequencies and their corresponding response probabilities was *r* = .653, these findings suggest that differences in response probabilities do not necessarily produce significant differences in response frequencies. Moreover, the goal of increasing response frequencies may not be of any significance or importance given Kapur's (2011) finding that response probabilities (not frequency counts) observed in students' discourse is what predicts learning outcomes achieved in groups.

Overall comparison of message-response patterns

The state diagrams in Fig. 6 reveal more or less the same similarities, differences, and trends reported previously in the analysis of the state diagrams in Fig. 4. The main finding of particular interest was produced by cross comparing the state diagrams in Fig. 6. A cross comparison revealed that when high-GSE students engaged in exchanges with other high-GSE students and when low-GSE students engaged in exchanges with low-GSE students, the differences in response rates to messages posted by high- versus low-GSE students noted previously in Fig. 4 were further amplified. In Fig. 4, the transitional probabilities between ARGL → BUT (.41) and ARGH → BUT (.58) differed by .17. This difference increased to .19 when examining the transitional probabilities in Fig. 6 between ARGL → BUTL (.34)

and ARGH → BUTH (.53). Likewise, the difference of .08 between BUTL → BUT (.62) and BUTH → BUT (.70) increased to .17 between BUTL → BUTL (.58) and BUTH → BUTH (.75). The difference of .01 between BUTL → EVID (.13) and BUTH → EVID (.12) increased to .06 between BUTL → EVIDL (.16) and BUTH → E-VIDH (.10). The difference of .06 between BUTL → EXPH (.24) and BUTH → EXPH (.18) increased to .11 between BUTL → EXPLL (.26) and BUTH → EXPLH (.15).

This set of findings suggests that students' response to messages may have been influenced not only by the GSE of the students (and their perceived credibility) that posted the message, but also by the grammar/spelling skills and perceived credibility of the student responding to messages. One plausible explanation is that the perceived credibility/authority of one student is lower from a relative standpoint of another student who perceives him/herself to be high or higher in credibility/authority. In contrast, the perceived credibility of the one student may be relatively higher from the standpoint of a student who perceives him/herself to be low or lower in credibility/authority. The implication of this finding is that the level of argumentation in online debates could be influenced by the group's composition in terms of students' skills in grammar and spelling.

In Fig. 2, the multitude of probabilities that were significantly *higher* than expected in the upper left quadrant of the transitional probability matrix (combined with the multitude of probabilities that were significantly *lower* than expected in the lower left quadrant) suggests that high-GSE students were more likely to post replies to other high-GSE students than to low-GSE students. The multitude of probabilities that were significantly *higher* than expected in the lower right quadrant of the transitional probability matrix (combined with the multitude of probabilities that were significantly *lower* than expected in the upper right quadrant) suggests that the low-GSE students were more likely to post replies to other low-GSE students than to high-GSE students. This pattern suggests that students may be more inclined (or perhaps more comfortable) to engage in argumentative exchanges with students that exhibit similar levels of GSE and/or similar level of perceived credibility/authority.

Response patterns in exchanges between students across groups

Data analysis

To determine the potential effects of GSE in *heterogeneous* groups, post hoc analysis was conducted to test for differences in response probabilities produced in: a) the responses of high-GSE students to the messages from low-GSE students; and b) the responses of low-GSE students to the messages from high-GSE students. The purpose of this analysis was to

	ARGH	BUTH	EVIDH	EXPLH	Replies	No Replies	Givens	Reply Rate		ARGL	BUTL	EVIDL	EXPLL	Replies	No Replies	Givens	Reply Rate
ARGH	.03	**.53**	.16	.28	133	22	112	.80	ARGL	.01	**.34**	**.35**	.30	168	22	124	.82
	1.94	-2.84	1.27	1.71						1.36	-5.75	5.09	1.60				
BUTH	.00	**.75**	.10	**.15**	134	172	322	.47	BUTL	.00	**.58**	**.16**	.26	180	211	393	.46
	-1.78	3.85	-1.40	-2.81						-.78	2.23	-2.65	.04				
EVIDH	.00	.64	.19	.17	36	38	73	.48	EVIDL	.00	.59	.24	.17	58	79	143	.45
	-.76	.15	1.24	-.94						-.37	1.11	.42	-1.62				
EXPLH	.02	.52	.09	**.37**	46	54	107	.50	EXPLL	.00	**.72**	**.06**	.23	71	115	190	.39
	.45	-1.59	-.91	2.43						-.42	3.66	-3.61	-.72				
	5	219	45	80	349	738	1557	.32		1	247	105	124	477	738	1557	.32

Fig. 5 Screen shot from DAT with aggregated response frequencies and probabilities to identify patterns in responses to messages exchanged between high-GSE and between low-GSE students
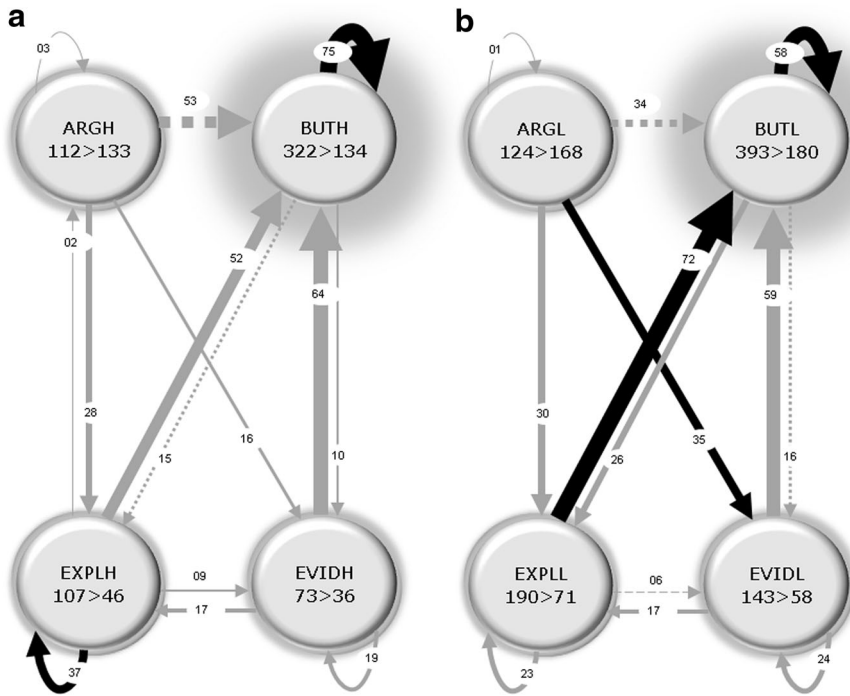
Fig. 6 Transitional state diagrams revealing patterns in message-response exchanges between high-GSE students and between low-GSE students. **a** Message-response exchanges between students with HIGH error. **b** Message-response exchanges between students with LOW error

determine if level of argumentation can be improved by forming groups that are homogeneous or heterogeneous in GSE by comparing the *highest* of these response probabilities to the *lowest* probabilities observed in the exchanges between students of similar skill. The left table in Fig. 7 presents the probabilities of responses from low-GSE student to the messages of high-GSE students, and the right table presents the probabilities of responses from high-GSE students to the messages of low-GSE students.

Differences in response probabilities

Arguments in the exchanges between low-GSE students elicited proportionately fewer challenges (.34) than in the exchanges where low-GSE students replied with challenges to arguments from high-GSE students (.73), ø = .802; *SE* = .04. Given that a strong association was found between this response pattern and group composition, the finding suggests that heterogeneous groups can nearly double the likelihood of producing challenges in reply to arguments than groups with only low-GSE students. Challenges in exchanges between low-GSE students elicited proportionately fewer counter-challenges (.58) than in the exchanges where high-GSE students replied with challenges to challenges from low-GSE students (.75), ø = .184; *SE* = .102. Challenges in exchanges between high-GSE students elicited proportionately fewer responses with supporting evidence (.10) than in the exchanges where low-GSE students replied with supporting evidence to the challenges of high-GSE students (.17), ø = −.358; *SE* = .189. Challenges in exchanges between high-

Low-Error Replies to High-Error Students

	ARGL	BUTL	EVIDL	EXPLL	Replies
ARGH	.02	.73	.11	.13	45
	1.40	.78	-.61	-.74	
BUTH	.00	**.59**	.17	.24	59
	-.90	-2.15	1.00	1.96	
EVIDH	.00	.78	.11	.11	9
	-.27	.59	-.23	-.46	
EXPLH	.00	.84	.11	.05	19
	-.41	1.55	-.43	-1.44	
	1	91	18	22	132

High-Error Replies to Low-Error Students

	ARGH	BUTH	EVIDH	EXPLH	Replies
ARGL	**.07**	.61	.19	.13	54
	2.68	-.72	1.14	-1.26	
BUTL	.00	**.75**	**.05**	.19	57
	-1.60	2.13	-2.46	.26	
EVIDL	.00	.65	.29	.06	17
	-.73	-.01	1.91	-1.40	
EXPLL	.00	**.45**	.15	**.40**	20
	-.80	-2.00	.11	2.71	
	4	96	21	27	148

Fig. 7 Screen shot from DAT with aggregated response frequencies and probabilities to examine patterns in low-GSE students' replies to high-GSE students and vice versa

GSE students elicited proportionately fewer explanations (.15) than in the exchanges where low-GSE students replied with explanations to the challenges of high-GSE students (.24), $\emptyset = -.322$; $SE = .163$. These observed differences in response probabilities (and the associations found between errors and group composition) suggest that higher levels of argumentation can be achieved by placing students into groups that are heterogeneous (rather than homogeneous) in grammar/spelling skill.

Association between use of qualifiers and GSE

One of the underlying assumptions of this study was that GSE can affect response patterns because GSE can negatively impact one's perceived level of credibility. Similarly, a prior study (Jeong, 2005a) found that the use of qualifiers (e.g., maybe, might, probably, I think, though) affected how other's respond to one's messages (based on the assumption that one's use of qualifiers can also lower one's perceived credibility). As a result, additional analysis was conducted to test the relationship between the number of qualifiers ($M = 1.00$, $STD = .871$, $n = 453$) and number of GSE ($M = .494$, $STD = .780$, $n = 453$) observed within each message. After removing 12 messages with outlying values in number of errors and/or qualifiers at 2 standard deviations and above, a significant correlation was found between number of qualifiers and errors, Spearman rank correlation $R = .348$, $p = .000$, $n = 453$. In addition, the mean number of GSE in messages with qualifiers ($M = 1.179$, $STD = .846$, $n = 452$) was significantly greater than the number of errors in messages with no qualifiers ($M = .909$, $STD = .870$, $n = 452$), $t(452) = -.317$, $p = .001$. One plausible reason for the observed association between GSE and use of qualifiers is that the students that produced more errors may be students that were less competent than the students that produced fewer GSE, and hence were less confident and certain about the accuracy and veracity of the claims they presented in the debate.

Discussion

Given that no prior studies have examined how GSE affect message-response exchange in online discussions, this study produced the first findings to show that GSE can affect the nature of student-to-student interactions and level of argumentation in online debates. The

main findings in this study were the following: a) arguments posted by high-GSE students elicited significantly more challenges than arguments posted by low-GSE students, while the challenges posted by high-GSE students (though not statistically significant) showed the tendency to elicit fewer replies with counter-challenges, supporting evidence, and explanations; b) these differences in response probabilities were magnified when high-GSE students responded with challenges to the arguments of other high-GSE students; and c) level of argumentation was higher when students of different levels of error exchanged messages and responses with one another compared to exchanges between only high-GSE students and/or only low-GSE students. Overall, responding to arguments with a challenge was the one type of exchange most likely to be affected by students' GSE. The other three types of exchanges (challenge → counter-challenge, challenge → supporting evidence, challenge → explanation) were also found to be affected by GSE, but its effects were very small. These negative yet small effects are consistent with the findings in Carpenter's (2012) study which found the effects of verbal stumbles on persuasiveness and message acceptance to be negative but small.

The main explanation for these findings can be drawn from prior studies showing that: (1) GSE in written communications can affect perceived credibility (Liu, 2004; Morris et al., 2012) just like the way verbal fluency (e.g., pauses, stutters) in face-to-face communication can affect how people perceive the speaker's credibility and competence (Burgoon et al., 1990; Berger, 1997; Myers, and Bryant, 2004); and most of all, (2) perceived credibility can influence the decision to accept or challenge the ideas presented by the speaker (Carpenter, 2012). Furthermore, the students in this study that made the most GSE showed a moderate tendency to use more qualifiers to hedge their claims. These two factors together (GSE and use of qualifiers) help to explain why the arguments posted by high-GSE students were less likely to be accepted by other students and instead were more likely to be challenged by students than the arguments posted by low-GSE students (students that used fewer qualifiers and were perceived to be more competent and credible). The same explanation can be applied to the differences in the probability of responses to challenges with counter-challenges, supporting evidence, and explanations. One reason as to why the effects of GSE on the responses to challenges (counter-challenge, supporting evidence, explanations) were smaller than its effects on the response to arguments with challenges is that these three particular types of exchanges occurred later in a discussion thread following the posting of an argument. Previous studies have found that the later a message is posted to the discussion thread, the less likely the message is to elicit a response due to the growing number of new messages posted to the discussion forum that competes for students' attention and responses (Hewitt, 2005; Jeong, 2004). As a result, the probability of a response of any kind to a given challenge was likely depressed by this inherent nature of online threaded discussions.

In regards to the instructional implications, the findings suggest that instructors encourage students to: (1) minimize GSE by using a spell checker; and (2) do their best to focus less attention on GSE while focusing more attention on the substantive nature (e.g., accuracy, veracity, relevance) of other students' postings. By doing so, students may be more likely to evaluate and rebut all presented arguments (and not just the arguments of the select few) and as a result, exchange and integrate more diverse viewpoints from all participants. In addition, the findings suggest that instructors take the necessary steps to form discussion groups that are heterogeneous in GSE (or more broadly speaking, heterogeneous in writing skills, language proficiency, and/or nationality) in order to maximize opportunities for all students to engage in high-level argumentation. These practices can help to ensure that all ideas posted by all students regardless of their writing

skills are critically discussed and cross-examined to evaluate and establish the accuracy and veracity of each claim presented in a debate. Future studies down the line will be needed to determine (using controlled experiments) whether these strategies promote higher levels of argumentation, and to what extent improvements in argumentation help to achieve particular outcomes like better group decision making and problem solving. Such tested and proven strategies will become increasingly important as international partnerships in distance learning programs grow in number and bring together students with more diverse language and writing proficiencies.

However, the findings in this study are not necessarily conclusive. Future studies will need to: (a) measure the quality of students' postings (e.g., accuracy, clarity, relevance, veracity) and its effect on how students respond to the posting to isolate and determine the relative effects of students' GSE and use of qualifiers; (b) control for and test the effects of group composition; (c) control for GSE by comparing two debates—one with and one without the use of grammar/spelling checking—to verify the cause-effect relationship between GSE and level of argumentation; (d) directly measure how students' perceptions of credibility/authority are affected by students' GSE to determine to what extent response patterns are affected by GSE versus perceptions of credibility versus message quality; (e) test the effects of GSE observed in each type of message when messages are posted anonymously; (f) examine exchanges produced by students with greater variance in GSE (e.g., undergraduates, high school students) to seek out larger differences in response patterns and test the effects of the observed patterns on specific learning outcomes; (g) control for prior face-to-face contact between students; and (h) omit the use of message labels in students' postings and remove minimum number of posting requirements to increase the generalizability of the findings (although it is not known as to whether message labels affect responses to messages with errors more than messages without errors, or vice versa). All of these noted limitations provide directions for further research to achieve a better understanding of the effects of students' grammar/spelling skills (or other associated characteristics of students' messages) on level of critical discourse in online discussions.

References

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

Baker, M. (1999). Argumentation and constructive interaction. In P. Courier & J. E. B. Andriesseni (Eds.), *Foundations of argumentative text processing* (pp. 179–202). Amsterdam: Amsterdam University Press.

Baker, M., & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted learning, 13*, 175–193.

Bakhtin, M. (1981). Discourse in the novel (M. Holquist & C. Emerson, Trans.). In M. Holquist (Ed.), *The dialogic imagination* (pp. 259–422). Austin: University of Texas Press.

Beers, P.J., Boshuizen, E., & Kirschner P. (2004). Computer Support for knowledge construction in collaborative learning environments. Paper presented at the American Educational Research Association Conference, San Diego, CA.

Berger, C. (1997). *Planning Strategic Interaction: Attaining Goals through Communicative Action* (p. 138). Mahwah, NJ: Lawrence Erlbaum.

Burgoon, J. K., Birk, T., & Pfau, M. (1990). Nonverbal behaviors, persuasion, and credibility. *Human Communication Research, 17*, 140–169.

Carpenter, C. (2012). A meta-analysis and an experiment investigating the effects of speaker disfluency on persuasion. *Western Journal of Communication, 76*(5), 552–569.

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

Duffy, T. M., Dueber, B., & Hawley, C. L. (1998). Critical thinking in a distributed environment: A pedagogical base for the design of conferencing systems. In C. J. Bonk & K. S. King (Eds.), *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse* (pp. 51–78). Mahwah, NJ: Erlbaum.

Ertmer, P. A., Richardson, J. C., Belland, B., Camin, D., Connolly, P., Coulthard, G., et al. (2007). Using peer feedback to enhance the quality of student online postings: An exploratory study. *Journal of Computer-Mediated Communication, 12*(2), 412–433.

Garrison, D. (2007). Online community of inquiry review: Social, cognitive, and teaching presence issues. *Journal of Asynchronous Learning Networks, 11*(1), 61–72.

Garrison, D. R., & Archer, A. (2003). Community of inquiry framework for online learning. In M. Moore (Ed.), *Handbook of distance education*. New York: Erlbaum.

Hew, K. F., Cheung, W. S., & Ng, C. S. L. (2010). Student contribution in asynchronous online discussion: a review of the research and empirical exploration. *Instructional Science, 38*(6), 571–606.

Hewitt, J. (2005). Toward an understanding of how threads die in asynchronous computer conferences. *The Journal of Learning Sciences, 14*(4), 567–589.

Jeong, A. C. (2004). 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.

Jeong, A. C. (2005a). The effects of linguistic qualifiers on group interaction patterns in computer-supported collaborative argumentation. *International Review of Research in Open and Distance Learning, 6*(3). Retrieved November 3, 2005, from http://www.irrodl.org/index.php/irrodl/article/view/258/401.

Jeong, A. C. (2005b). A guide to analyzing message-response sequences and group interaction patterns in computer-mediated communication. *Distance Education, 26*(3), 367–383.

Jeong, A. C. (2006). The effects of conversational styles of communication on group interaction patterns and argumentation in online discussions. *Instructional Science, 34*(5), 367–397.

Jeong, A. C. (2013). Discussion analysis tool. Retrieved August 12, 2005, from http://myweb.fsu.edu/ajeong.

Johnson, D., & Johnson, R. (2009). Energizing learning: The instructional power of conflict. *Educational Researcher, 38*(1), 37–51.

Jonassen, D. H., & Kwon, H. I. (2001). Communications patterns in computer mediated versus face-to-face group problem solving. *Educational Technology Research and Development, 49*(1), 35–51.

Kapur, M. (2011). Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of Computer-Supported Collaborative Learning, 6*(1), 39–56.

Karacapilidis, N., & Papadias, D. (2001). Computer supported argumentation and collaborative decision making: The Hermes system. *Information Systems, 26*(4), 259–277.

Koschmann, T. (1999). Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration. In C. Hoadley & J. Roschelle (Eds.), *Proceedings of the computer support for collaborative learning (CSCL) 1999 conference*, December 12–15, Stanford University, Palo Alto, California. Mahwah, NJ: Lawrence Erlbaum Associates.

Leinonen, T., Virtanen, O., & Hakkarainen, K. (2002). Collaborative discovering of key ideas in knowledge building. In G. Stahl (Ed.), *Proceeding of the conference on computer support for collaborative learning: foundations for a CSCL Community*. Boulder, CO: CSCL '02

Lemus, D., Seibold, D., Flanagin, A., & Metzger, M. (2004). Argument and decision making in computer-mediated groups. *Journal of Communication, 54*(2), 302–320.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Liu, Z. (2004). Perceptions of credibility of scholarly information on the web. *Information Processing and Management, 40*(6), 1027–1038.

McAlister, S., Ravenscroft, A., & Scanlon, E. (2004). Combining interaction and context design to support collaborative argumentation using a tool for synchronous CMC. *Journal of Computer Assisted Learning: Special Issue: Developing Dialogue for Learning, 20*(3), 194–204.

Meyer, K. (2003). Face-to-face versus threaded discussion: The role of time and higher-order thinking. *Journal of Asynchronous Learning Networks, 7*(3), 55–65.

Morris, M., Counts, S., Roseway, A., Hoff, A., & Schwarz, J. (2012). Tweeting is believing? Understanding microblog credibility perceptions. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, February 11–15, 2012. Seattle, Washington. New York, NY: ACM. (pp. 441–450)

Myers, S., & Bryant, L. (2004). College students' perceptions of how instructors convey credibility. *Qualitative Research Reports in Communication, 5*, 22–27.

Noroozi, O., Weinberger, A., Biemans, H., Mulder, M., & Chizari, M. (2012). Argumentation-based computer supported collaborative learning (ABCSCL): a synthesis of 15 years of research. *Educational Research Review, 7*(2), 79–106.

Pena-Shaff, J., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in Computer Bulletin Board (BBS) discussions. *Computers & Education, 42*, 243–265.

Ravenscroft, A., Sagar, M., Baur, E., & Oriogun, P. (2008). Ambient pedagogies, meaningful learning and social software. In S. Hatzipanagos & S. Warburton (Eds.), *Social software & developing community ontologies* (pp. 432–450). Hershey, PA: IGI Global Publishing.

Rourke, L., & Kanuka, H. (2007). Computer conferencing and distance learning. In H. Bidgoli (Ed.), *The handbook of computer networks* (Vol. 3, pp. 831–842). Hoboken, NJ: Wiley.

Scardamalia, M., & Bereiter, C. (1996). Computer support for knowledge-building communities. In T. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm* (pp. 249–268). Mahwah NJ: Erlbaum.

Schwarz, B. B., & Glassner, A. (2007). The role of floor control and of ontology in argumentative activities with discussion-based tools. *International Journal of Computer-Supported Collaborative Learning, 2*(4), 449–478.

Sloffer, S., Dueber, B., & 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. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*(4), 421–447.

Tannen, D. (1998). *The argument culture: Moving from debate to dialogue*. New York: Random House Trade.

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

Van Eemeren, F., Grootendorst, R., & Henkemans, F. (2008). Dialectical profiles and indicators of argumentative moves. *Journal of Pragmatics, 40*(3), 475–493.

Veerman, A., Andriessen, J., & Kanselaar, G. (2002). Collaborative argumentation in academic education. *Instructional Science, 30*(3), 155–186.

Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: An empirical investigation of a possible sociocultural model of cognitive development. *Learning and Instruction, 9*, 493–516.

Wiley, J., & Voss, J. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology, 91*, 301–311.

Dr. Allan Jeong is Associate Professor in Instructional Systems at Florida State University. His research is focused on developing software tools and methods to measure, visualize, and model computer-supported collaborative learning processes.

Haiying Li is an experimental psychology Ph.D. student working in the Institute for Intelligent Systems at the University of Memphis. She is interested in discourse analysis, intelligent tutoring systems, motivation, and computational linguistics.

Andy Jiaren Pan is a graduate of the Instructional Systems Master's program at Florida State University with professional interests in instructional design and development.