Routledge
Taylor & Francis Group

Interactive Learning Environments

ISSN: 1049-4820 (Print) 1744-5191 (Online) Journal homepage: https://www.tandfonline.com/loi/nile20

Students' online argumentative peer feedback, essay writing, and content learning: does gender matter?

Omid Noroozi, Javad Hatami, Arash Bayat, Stan van Ginkel, Harm J. A. Biemans & Martin Mulder

To cite this article: Omid Noroozi, Javad Hatami, Arash Bayat, Stan van Ginkel, Harm J. A. Biemans & Martin Mulder (2018): Students' online argumentative peer feedback, essay writing, and content learning: does gender matter?, Interactive Learning Environments, DOI: 10.1080/10494820.2018.1543200

To link to this article: https://doi.org/10.1080/10494820.2018.1543200

Published online: 20 Nov 2018.

INTERACTIVE LEARNING ENVIRONMENTS
https://doi.org/10.1080/10494820.2018.1543200

Routledge
Taylor & Francis Group

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

Students' online argumentative peer feedback, essay writing, and content learning: does gender matter?

Omid Noroozi [a,b], Javad Hatami[a], Arash Bayat[c], Stan van Ginkel[b,d], Harm J. A. Biemans [b] and Martin Mulder [b]

[a]Tarbiat Modares University, Tehran, Iran; [b]Wageningen University, Wageningen, Netherlands; [c]Islamic Azad University of Sanandaj, Sanandaj, Iran; [d]Archimedes Institute, University of Applied Sciences, Utrecht, Netherlands

ABSTRACT

Whilst the importance of online peer feedback and writing argumentative essays for students in higher education is unquestionable, there is a need for further research into whether and the extent to which female and male students differ with regard to their argumentative feedback, essay writing, and content learning in online settings. The current study used a pre-test, post-test design to explore the extent to which female and male students differ regarding their argumentative feedback quality, essay writing and content learning in an online environment. Participants were 201 BSc biotechnology students who wrote an argumentative essay, engaged in argumentative peer feedback with learning partners in the form of triads and finally revised their original argumentative essay. The findings revealed differences between females and males in terms of the quality of their argumentative feedback. Female students provided higher-quality argumentative feedback than male students. Although all students improved their argumentative essay quality and also knowledge content from pre-test to post-test, these improvements were not significantly different between females and males. Explanations for these findings and recommendations are provided.

ARTICLE HISTORY
Received 1 October 2017
Accepted 1 September 2018

KEYWORDS
Argumentative essay; online peer feedback; gender; learning; writing

Introduction

Argumentative essay writing is considered as a key educational objective and a popular activity for higher education students (see Asterhan, 2018; Noroozi et al., 2016; Wu, 2006), especially when they deal with complex and controversial issues. Argumentation has always been regarded as a crucial component of each essay (Wingate, 2012). In social constructivist learning paradigms, learners engage in discussions with learning peers, argue and negotiate meaning with them in order to learn about the topic, (co) construct knowledge and/or solve complex issues (see Noroozi et al., 2012). Based on adjusted Toulmin's (1958) model of argumentation, writing an argumentative essay on controversial topics requires a clear position on the issue at hand that should be followed by logical evidence and arguments to support that particular position. Also, such essays require expression of the opposing views in the forms of counter-arguments that of course need to be refuted. As a result, one must integrate various pros and cons of the issue at stake while taking into account the opinions and perspectives of both the advocates and the opponents followed by a clear conclusion on the topic (see Noroozi et al., 2016; Andrews, 1995; Qin & Karabacak, 2010; Toulmin, 1958; Wood, 2001).

CONTACT Omid Noroozi omid.noroozi@wur.nl

Writing high-quality argumentative essays is not an easy task for higher education students. Teachers are typically not satisfied with regard to the overall argumentative quality of students' essays (see Noroozi et al., 2016; Graham & Perin, 2007; Pessoa, Mitchell, & Miller, 2017) and thus complain that most students have difficulties with essay writing (Kellogg & Whiteford, 2009). That is why various instructional strategies and practices have been used to enhance students' writing capacities (Stern & Solomon, 2006). Peer feedback, among other things, is a prominent approach that is used to enhance students' writing as well as content learning, and motivation in the particular domain (Nelson & Schunn, 2009; Van Ginkel, Gulikers, Biemans, & Mulder, 2015). Peer feedback has the potential to enable students to find the gap between their own current and ideal states resulting on how and what to improve (see Lizzio & Wilson, 2008).

With the advancement of educational technologies, online learning environments can be designed for the realization of peer feedback which has both informational value for supporting writing and learning, and motivational value for stimulating students' efforts (Shute, 2008). The benefit of online peer feedback over traditional peer feedback is that students are able to present and submit their contributions and also re-review learning partners' submissions in a more structured way, without restriction (Lin, Liu, & Yuan, 2001). Online peer feedback settings allow for the implementation of various types of scripts and scaffolding peer feedback processes that can guide learners towards a desirable mode of and argumentation interaction (Noroozi et al., 2016). Furthermore, it provides students with the flexibility to modify their feedback through the learning processes against face-to-face and paper-based feedback settings that offer less opportunities for the validity and reliability of peer feedback (Noroozi et al., 2016; Mostert & Snowball, 2013; Yang, 2011).

Argumentative peer feedback provides learners with the opportunity to critically test, enlighten, and analyse learning partners' arguments and understand multiple perspectives of the issues at stake that can, in turn, lead to writing high-quality argumentative essays (Noroozi et al., 2016). Furthermore, giving and receiving feedback on one another's argumentative essays would contribute to content learning since deepening and broadening the debate with detailed justifications and elaborations may lead to knowledge construction and content learning (Noroozi et al., 2016; Munneke, Andriessen, Kanselaar, & Kirschner, 2007; Van Amelsvoort, Andriessen, & Kanselaar, 2007).

The structure of argumentative essays and also argumentative peer feedback can be linked to socio-constructivist and socio-cognitive theory (Coffin & O'Halloran, 2008). Argumentative peer feedback, from this perspective, can be seen as part of an interactive process between learners with peers or experts. Such feedback followed by reasoned debate is argued to be central to the process by which higher-order mental thinking, critical reasoning, and reflection are developed (McAlister et al., 2004). Finally, this study follows the basic principles of Piaget's theory of cognitive development and its implications for peer learning. This theory explains how peer learning relates to Piaget's model of constructivism in terms of learning and cognitive change through discussions, logical thinking, and reasoning with peers and why it is more effective than independent and individual learning (see De Lisi & Golbeck, 1999). Overall, based on the theories explained above, in active learning situations, students need to express their opinions in the form of writing argumentative essays on what they are thinking, what they are believing, what they are meaning, and what they are promoting with regard to controversial issues in their own disciplines (see Noroozi et al., 2016).

There is a massive scientific research in terms of the importance of students' essay writing and the impacts of peer feedback on essay writing and learning (see Noroozi et al., 2016; Bayerlein, 2014; Gabelica, Van den Bossche, Segers, & & Gijselaers, 2012). However, an issue that is left under investigated is the extent to which female and male students differ in terms of their argumentative essay writing. Most studies on gender and argumentation have been related to the argumentative discourse rather than writing argumentative essays (e.g. Asterhan, Schwarz, & Gil, 2012; Erkens & Janssen, 2008). For instance, in a study by Prinsen, Volman, Terwel, and Van den Eeden (2009), it

was found that males elaborate on their messages less than females during peer feedback processes. Furthermore, males disagreed with their learning partners more often than females during argumentative peer feedback (see also Selfe & Meyer, 1991). Caspi, Chajut, and Sapporta (2008) reported no difference between females and males in online classroom discussions in terms of qualitative features of the dialogue. Asterhan et al. (2012) showed that female groups score higher than males on argumentation quality such as the inclusion of complex arguments and alternative standpoints. Erkens and Janssen (2008) found that females tend to apply more affiliative language such as argumentative and responsive dialogue acts against males who use more assertive language such as imperative and informative dialogue acts. Underwood, Underwood, and Wood (2000) found that males use more authoritative propositions and claims while females use more statements on their own intuitive opinions and personal conceptions. Li (2002) revealed that female' initial messages include a fewer explanation on the issue at stake than males' initial messages. Prinsen, Volman, and Terwel (2007) showed that disagreement is more seen in male communication and writing styles. Furthermore, they found that males use more authoritative statements than females while females provide less explanations than males.

First, although the role of gender for argumentative discourse is well studied (see above), yet little studies have addressed to what extent gender plays a role for the way students engage in argumentative peer feedback processes. So far, most research studies on gender and argumentation have been related to the argumentative discourse quality and not the quality of argumentative feedback in online peer feedback settings. It is not clear how female and male students provide argumentative feedback on their learning partners' essays during peer feedback processes. Researchers and educational designers need to understand how the quality of argumentative feedback of female and male students differs in an online peer feedback setting because this could have consequences for the way learning groups can be composed. Therefore, the first aim is to explore the extent to which female and male students differ in terms of their argumentative feedback quality.

Second, little studies have addressed the extent to which gender plays a role for the way students write argumentative essays in online peer feedback settings. Previous studies have not yet explored whether female and male students write argumentative essays of different quality. Furthermore, it is unclear to what extent female and male students respond differently to the argumentative feedback of their learning partners in relation to their revised written argumentative essays. Therefore, the second goal is to explore the extent to which female and male students differ in terms of their written argumentative essays. Furthermore, the aim is to explore the extent to which female and male students respond differently to the argumentative feedback of learning partners, as shown in their revised written argumentative essays.

And finally, it is important to find out to what extent female and male students benefit differently from the argumentative feedback of their learning partners in relation to their content learning. The reason is that with argumentative peer feedback, students are given the opportunity to reflect on the materials and the content, and to broaden and deepen their understanding on the issues at stake because they can compare their own contributions with other learning partners (Yang, 2010).

Research questions

Clearly, more research is needed to investigate the role of gender and engendered differences in students' argumentative feedback quality, written argumentative essays, and content learning. Following research questions address the goals of this study:

1. To what extent female and male students differ regarding argumentative feedback quality in online peer feedback settings?
2. To what extent female and male students differ regarding writing argumentative essays and their responses to the argumentative feedback of learning partners in online peer feedback settings?

3. To what extent female and male students benefit differently from the argumentative feedback of their learning partners in relation to their content learning in online peer feedback settings?

Method

Overall, 201 undergraduate students, who enrolled for the course "Introduction Molecular Life Sciences and Biotechnology", from a Dutch university in the domain of life sciences participated in this study. However, for data analyses, 189 students (63 triads) were included in the study based on their completion of all tasks. The average age of the participants was 19.20 years. Students were randomly assigned into triads who were distributed over different classrooms. The topic was the use of "cultured meat manufacturing – insect cells" as part of the bigger concept of the Genetically Modified Organisms (GMOs). All activities were conducted using a self-designed digital learning module that presents information in various forms, e.g. diagrams, texts, and pictures (see Figure 1). Furthermore, this module offers the context that is used for a user-friendly interaction style for reasoned and structured feedback as well as justified and logical arguments. Various input text boxes are used in this module for promoting students' critical thinking, argumentation, reasoning and argumentative processes and activities (see Figure 2).

The study took about 240 minutes comprising of four phases. In the first phase, students received a verbal introduction to the module during class time (10 minutes). Then, they completed several surveys on demographic variables e.g. age, gender, study programme, and student's opinion on GMOs followed by a test on their prior content knowledge on the topic (20 minutes). In the second phase, students were given the opportunity to read theoretical materials, texts, and scientific publications on the topic (40 minutes), and to write an essay individually (30 minutes) for almost 500 words on this statement: "Insect-cell biomass infected with genetically modified baculovirus is a

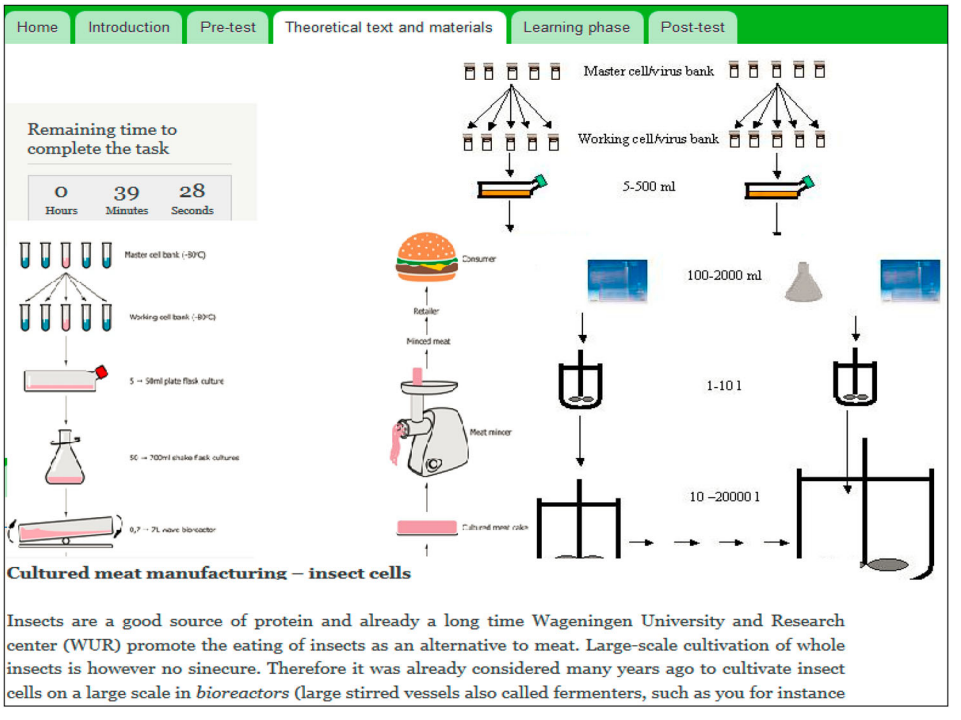

Figure 1. The online learning environment: In this phase, students are asked to read the theoretical text and materials in order to be prepared for writing individual argumentative essay.

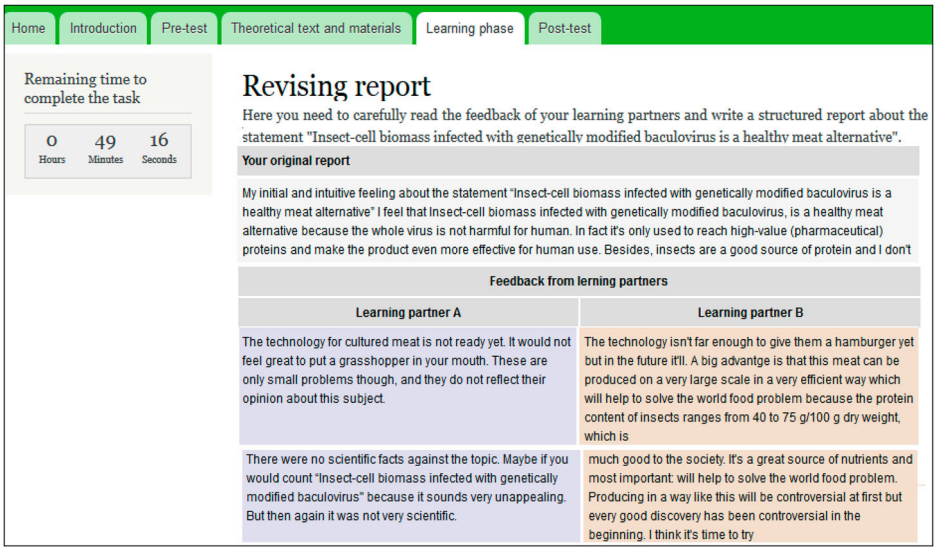

Figure 2. The online learning environment: In this phase, students see their own original essay, read the feedback of the two learning partners, and finally revise their original essay based on the feedback.

healthy meat alternative''. In the third phase, every student was asked to carefully read the essays of the two other partners in her/his group and provide specific feedback on them for about 50 minutes (25 minutes per each essay). Finally, the post-test phase took place (80 minutes). In this phase, students were first asked to carefully read the detailed feedback of the other two partners (25 minutes) and then write a revised essay for about 500 words (30 minutes). Students then completed several questionnaires on content knowledge and opinion on GMOs' issues (20 minutes).

The coding scheme designed by Noroozi et al. (2016) was used to measure the quality of argumentative essays and feedback of the students both in the pre-test, post-test and during feedback process. This instrument was designed based on the literature (see Andrews, 1995; Qin & Karabacak, 2010; Toulmin, 1958; Wood, 2001) and later adjusted to biotechnology domain (see Noroozi et al., 2016). This coding scheme comprised of a list of variables (i.e. Intuitive opinion, Pro-claims, Justification for pro-claims, Con-claims, Justification for con-claims, Integration of pros and cons, Conclusion) that reflect the quality of written argumentative essays of students and labelled according to Table 1. For all students, a single score was assigned for each of these variables ranging from zero to two. The reliability index of inter-rater agreement with respect to the quality of the argumentative essays between the two trained coders for about 10% of the data resulted in identical scores in 98% and 91% of the contributions in the pre- and post-test respectively. The inter-rater agreement with respect to the argumentative feedback between the two coders for about 10% of the data resulted in identical scores in 85% of the data.

For this study, a survey was used to measure students' content learning both in the pre-test and post-test. This survey included 17 multiple-choice questions. The validity of this instrument had been obtained earlier through a panel of experts and the teachers of the course by organizing a series of expert consultation meetings. After several sessions of expert consultations, the experts and teachers had reached consensus on the 17 multiple-choice questions that matched well with the context of this study in the field of biotechnology. This survey has been used with high validity and reliability in several other studies such as Noroozi & Mulder (2017) and Noroozi et al. (2016). Every student was given one point for each correct answer resulting in a maximum of 17 points for both pre-test and post-test.

Table 1. Coding scheme to analyse the quality of students' argumentative essays and feedback with variables, points, labels and corresponding descriptions.

Variables	Points	Labels for argumentative essay quality	Descriptions of the labels for argumentative essay quality	Labels for feedback quality	Descriptions of the labels for feedback quality
Intuitive opinion	Zero	Not mentioned	No intuitive opinion is presented.	Not mentioned	No feedback on intuitive opinion is presented.
	One	Non-elaborated intuitive opinion	Intuitive opinion is presented but it is not discussed or elaborated on or it is discussed in an insignificant way.	Non-elaborated feedback on intuitive opinion	Feedback on intuitive opinion is presented but justification for the feedback is not discussed or elaborated on or it is discussed in an insignificant way.
	Two	Elaborated intuitive opinion	Intuitive opinion is presented and it is discussed or elaborated on in a significant way.	Elaborated feedback on intuitive opinion	Feedback on intuitive opinion is presented and justification for the feedback is discussed or elaborated on in a significant way.
Claims in favour of the topic	Zero	Not mentioned	No claim is presented in favour of the topic.	Not mentioned	No feedback on pro-claim(s)is presented.
	One	Non-elaborated pro-claim	Only one claim is presented in favour of the topic.	Non-elaborated feedback on pro-claim(s)	Feedback on pro-claim(s) is presented but justification for the feedback is not discussed or elaborated on or it is discussed in an insignificant way.
	Two	Elaborated pro-claim	Multiple claims (two or more) are presented in favour of the topic.	Elaborated feedback on pro-claim(s)	Feedback on pro-claim(s) is presented and justification for the feedback is discussed or elaborated on in a significant way.
Justification for claim(s) in favour of the topic	Zero	Not mentioned	No justification for pro-claim(s) is presented. None of the pro-claim(s) are justified. Pro-claim(s) are just being presented without any back up/support in terms of presenting scientific facts, evidence, examples, figures etc.	Not mentioned	No feedback on scientific facts in favour of the topic is presented.
	One	Non-elaborated justification for pro-claim(s)	Justification for pro-claim(s) is presented. Pro-claim(s) are being backed up/supported (with scientific facts, evidence, examples, figures etc.) but they are not strongly connected to the pro-claim(s). Justification for pro-claim (s) are not discussed or elaborated on in a significant way.	Non-elaborated feedback on scientific facts	Feedback on scientific facts in favour of the topic is presented but justification for the feedback is not discussed or elaborated on or it is discussed in an insignificant way.
	Two	Elaborated justification for pro-claim(s)	Justification for pro-claim(s) is presented. Pro-claim(s) are being backed up/supported (with scientific facts, evidence, examples, figures etc.) and they are strongly connected to the pro-claim(s). Justification for pro-claim (s) are discussed or elaborated on in a significant way.	Elaborated feedback on scientific facts	Feedback on scientific facts in favour of the topic is presented and justification for the feedback is discussed or elaborated on in a significant way.
Claims against the topic	Zero	Not mentioned	No claim is presented against the topic.	Not mentioned	No feedback on con-claim(s)is presented.
	One	Non-elaborated con-claim	Only one claim is presented against the topic.	Non-elaborated feedback on con-claim(s)	Feedback on con-claim(s) is presented but justification for the feedback is not discussed or elaborated on or it is discussed in an insignificant way.

	Two	Elaborated con-claim	Multiple claims (two or more) are presented against the topic.	Elaborated feedback on con-claim(s)	Feedback on con-claim(s) is presented and justification for the feedback is discussed or elaborated on in a significant way.
Justification for claim(s) against the topic	Zero	Not mentioned	No justification for con-claim(s) is presented. None of the con-claim(s) are justified. Con-claim(s) are just being presented without any back up/support in terms of presenting scientific facts, evidence, examples, figures etc.	Not mentioned	No feedback on scientific facts against the topic is presented.
	One	Non-elaborated justification for con (s)-argument	Justification for con-claim(s) is presented. Con-claim(s) are being backed up/supported (with scientific facts, evidence, examples, figures etc.) but they are not strongly connected to the con-claim(s). Justification for con-claim (s) are not discussed or elaborated on in a significant way.	Non-elaborated feedback on scientific facts	Feedback on scientific facts against the topic is presented but justification for the feedback is not discussed or elaborated on or it is discussed in an insignificant way.
	Two	Elaborated justification for con(s)-argument	Justification for con-claim(s) is presented. Con-claim(s) are being backed up/supported (with scientific facts, evidence, examples, figures etc.) and they are strongly connected to the con-claim(s). Justification for con-claim (s) are discussed or elaborated on in a significant way.	Elaborated feedback on scientific facts	Feedback on scientific facts against the topic is presented and justification for the feedback is discussed or elaborated on in a significant way.
Integration of pros and cons	Zero	Not mentioned	No integration of pros and cons is presented.	Not mentioned	No feedback on integration of pros and cons is presented.
	One	Non-elaborated integration of pros and cons	Integration of pros and cons is presented but justification for the integration is not discussed or elaborated on or it is discussed in an insignificant way.	Non-elaborated feedback on integration of pros and cons	Feedback on integration of pros and cons is presented but justification for the integration is not discussed or elaborated on or it is discussed in an insignificant way.
	Two	Non-elaborated integration of pros and cons	Integration of pros and cons is presented and justification for the integration is discussed or elaborated on in a significant way.	Elaborated feedback on integration of pros and cons	Feedback on integration of pros and cons is presented and justification for the integration is discussed or elaborated on in a significant way.
Conclusion	Zero	Not mentioned	No conclusion is presented.	Not mentioned	No feedback on conclusion is presented.
	One	Non-elaborated conclusion	Conclusion is presented but it is not discussed or elaborated on or it is discussed in an insignificant way.	Non-elaborated feedback on conclusion	Feedback on conclusion is presented but justification for the feedback is not discussed or elaborated on or it is discussed in an insignificant way.
	Two	Elaborated conclusion	Conclusion is presented and it is discussed or elaborated on in a significant way.	Elaborated feedback on conclusion	Feedback on conclusion is presented and justification for the feedback is discussed or elaborated on in a significant way.

Table 2. Differences among female and male students in terms of mean scores for feedback quality (Max = 2; Min = 0).

Variables	Label	Feedback quality		Difference between males and females Statistics
		Mean	SD	
Intuitive opinion	Male	1.44	.48	$F(1, 187) = 5.43$, $p < .05$, $\eta^2 = .03$
	Female	1.32	.53	
	Total	1.26	.57	
Claims in favour of the topic	Male	1.24	.54	$F(1, 187) = .04$, $p = .85$
	Female	1.25	.56	
	Total	1.14	.53	
Justification for claim(s) in favour of the topic	Male	1.34	.48	$F(1, 187) = 6.10$, $p < .05$, $\eta^2 = .03$
	Female	1.21	.52	
	Total	1.22	.53	
Claims against the topic	Male	1.34	.48	$F(1, 187) = 2.18$, $p = .14$
	Female	1.26	.51	
	Total	0.98	.64	
Justification for claim(s) against the topic	Male	1.15	.52	$F(1, 187) = 3.75$, $p = .05$
	Female	1.04	.60	
	Total	1.10	.55	
Integration of pros and cons	Male	1.22	.53	$F(1, 187) = 2.23$, $p = .14$
	Female	1.14	.55	
	Total	1.10	.42	
Conclusion	Male	1.21	.35	$F(1, 187) = 3.70$, $p = .06$
	Female	1.14	.40	
	Total	1.44	.48	

Results

Research question 1

Results showed that female and male students significantly differ in terms of mean quality scores of their argumentative feedback, Wilks' $\lambda = .91$, $F(7, 181) = 2.52$, $p < .05$, $\eta^2 = .09$ (see Table 2). This difference was mainly due to the argumentative feedback quality of the intuitive opinion and justification for claim(s) in favour of the topic. Overall, female students produced a higher quality of feedback but only with respect to intuitive opinions and justification for pro-claims than male students (see Table 2).

Research question 2

In the pre-test, female and male students did not differ at all with respect to mean quality scores of their argumentative essay, Wilks' $\lambda = .97$, $F(7, 181) = .68$, $p = .69$ (see Table 3). In the post-test, there was a difference between female and male students with respect to mean quality scores of their argumentative essay, Wilks' $\lambda = .08$, $F(7, 181) = 2.09$, $p < .05$, $\eta^2 = .07$ (see Table 3). Overall, male students produced a higher quality of essays than females in the post-test but only with respect to the integration of the pros and cons.

Results showed that argumentative essay quality of both female and male students improved significantly from pre-test to post-test, Wilks' $\lambda = .55$, $F(7, 181) = 31.28$, $p < .01$, $\eta^2 = .55$ (see Table 3). These improvements were mostly related to the pro-claims, justification for pro-claims, claims against the topic, integration of pros and cons, and conclusion.

Results showed that female and male students responded differently to the argumentative feedback of the partners, in their revised argumentative essays, Wilks' $\lambda = .91$, $F(7, 181) = 2.43$, $p < .05$, $\eta^2 = .09$ (see Table 3). Overall, males performed better than females with respect to their responses to the argumentative feedback of the learning partners in their revised argumentative essays but only for the integration of pros and cons.

Results for research question 3

Results showed that, despite their gender, the knowledge of all students regarding the topic of the discussion improved significantly from pre-test to post-test, Wilks' $\lambda = .55$, $F(1, 187) = 162.90$, $p < .01$,

Table 3. Differences among female and male students in terms of mean scores for argumentative essays both in pre-test, post-test, and improvements from pre-test to post-test.

Variables	Difference between essay quality of males and females in the pre-test	Difference between essay quality of males and females in the post-test	Essay quality improvements of all students from pre-test to post-test	Difference between essay quality improvements of males and females from pre-test to post-test
Intuitive opinion	$F(1, 187) = .40$, $p = .53$	$F(1, 187) = .80$, $p = .37$	$F(1, 187) = .00$, $p = .96$	$F(1, 187) = 1.22$, $p = .27$
Claims in favour of the topic	$F(1, 187) = .10$, $p = .75$	$F(1, 187) = 2.36$, $p = .13$	$F(1, 187) = 4.11$, $p < .01$, $\eta^2 = .02$	$F(1, 187) = 2.17$, $p = .14$
Justification for claim(s) in favour of the topic	$F(1, 187) = .76$, $p = .38$	$F(1, 187) = .06$, $p = .81$	$F(1, 187) = 1.76$, $p < .01$, $\eta^2 = .01$	$F(1, 187) = .31$, $p = .58$
Claims against the topic	$F(1, 187) = .02$, $p = .88$	$F(1, 187) = 1.27$, $p = .26$	$F(1, 187) = 55.93$, $p < .01$, $\eta^2 = .23$	$F(1, 187) = .46$, $p = .50$
Justification for claim(s) against the topic	$F(1, 187) = .42$, $p = .52$	$F(1, 187) = .09$, $p = .93$	$F(1, 187) = 3.21$, $p = .07$	$F(1, 187) = .17$, $p = .68$
Integration of pros and cons	$F(1, 187) = .78$, $p = .38$	$F(1, 187) = 9.66$, $p < .01$, $\eta^2 = .05$	$F(1, 187) = 86.96$, $p < .01$, $\eta^2 = .32$	$F(1, 187) = 8.52$, $p < .01$, $\eta^2 = .04$
Conclusion	$F(1, 187) = .49$, $p = .48$	$F(1, 187) = .99$, $p = .32$	$F(1, 187) = 50.51$, $p < .01$, $\eta 2 = .21$	$F(1, 187) = .12$, $p = .73$

$\eta^2 = .47$. However, the findings showed no significant difference between female and male students in terms of their improved content learning from pre-test to post-test, Wilks' $\lambda = .99$, $F(1, 187) = .15$, $p = .70$.

Discussions

Discussions of results for question 1

Female students constructed higher-quality argumentative feedback than males, especially when providing feedback on the intuitive opinion of the students regarding the GMOs and justification for pro-claims. This implies that females, compared with males, not only presented a detailed feedback on the intuitive opinions and also the arguments of their partners but also justified their feedback in a significant way. This finding is similar to the findings of Prinsen et al. (2009), concluding that females elaborate on their messages more than males during peer feedback processes. This finding also resembles previous literature suggesting that females typically score higher than males in terms of their argumentation quality such as elaborations on their arguments and considerations of multiple perspectives into account (see Asterhan et al., 2012). Furthermore, similar to the findings of this study, literature suggests that females more than males tend to apply affiliative language such as responsive and argumentative dialogue acts (Erkens & Janssen, 2008) and also focus on intuitive conceptions and personal opinions during argumentative discourse activities (Underwood et al., 2000).

Female and male students did not differ with respect to their argumentative feedback quality in terms of pro-claims, con-claims, justification for con-claims, integration of pros and cons, and conclusion. This is against the literature suggesting that males approach argumentative tasks differently than females (see Asterhan et al., 2012; Prinsen et al., 2009; Selfe & Meyer, 1991; Sullivan, Kapur, Madden, & Shipe, 2015). Based on the literature, we expected differences between males and females in their quality of written argumentative essays. The lack of such differences in the current study might be related to the cultural background of the participants. Scientific literature suggests that culture influences thinking, writing, and human behaviour (Hofstede, 1993) and hence each culture may have different patterns of argumentation and reasoning (Uysal, 2008, 2012). In this study, participants came from a western country, namely the Netherlands, which is considered as a small power distance society. In small power distance societies, students are encouraged to raise their opinion and discuss conflicts in their own knowledge beliefs regardless of their gender. In such cultures, the same educational goals and learning activities are expected from female and male students (Hsu, Van Dyke, & Smith, 2017). Thus, we speculate that the lack of differences in the quality of argumentative feedback between females and males for some variables of this study might be related to the equal distribution of the power distance between female and male students.

Discussions of results for question 2

The only difference between females and males in relation to their argumentative essay quality was that males received a higher score only in the post-test (and not the pre-test) and only with respect to the integration of pros and cons. The effect size was too small, though. So, when taking the effect size into account, females and males did not differ that much in terms of writing argumentative essays. Literature suggests that females not only during argumentative discourse but also when writing argumentative essays tend to elaborate more than males on their arguments, consider alternative perspectives, and express more intuitive conceptions and personal opinions (see Noroozi, 2018; Noroozi et al., 2012). This was not the case though based on the findings of this study both in the pre-test and the post-test. The plausible explanation could be related to the administration of the coding scheme in this study which ranged only between zero to two for each variable of the essay. This range was too small to reveal significant differences between males and females with a

large effect size. We speculate that if the range of the scoring was a bit wider (e.g. using a Likert scale ranging from zero to four instead of zero to two), we could have seen differences between males and females with regard to at least some variables of the argumentative essays as described in Table 1.

There was an improvement for the quality of almost all aspects of the argumentative essays of students from pre-test to post-test regardless of their gender. This implies that engaging in argumentative peer feedback processes can help students write higher-quality argumentative essays. Previous findings have shown the positive impacts of peer feedback strategies on students writing skills (see Gabelica et al., 2012; Kellogg & Whiteford, 2009). Literature suggests that giving and receiving argumentative feedback, elaborating on one's own and other's ideas, testing, enlightening, integrating multiple arguments, even opposing and disagreeing with the learning partners' arguments, and finally challenging others and being challenged by others can significantly contribute to the argumentation quality of the written essays (see Noroozi et al., 2016). Argumentative feedback reception from and provision for the other learning partners in the group along with elaborations and justifications helped students detect the gap between their own current argumentative quality level and the expected level. Making comparison between their own essays with others on what and how to improve somewhat provoked students' deep reflection and thinking that in turn was reflected in their post-test argumentative essays (see De Nisi & Kluger, 2000).

The improvement in the quality of the students' argumentative essay from pre-test to post-test was only significantly different with regard to the integration of the pros and cons favouring males than females. The reason could be that typically females tend to write more explanations and longer messages than males (see Prinsen et al., 2009). The order of the seven variables of the argumentative essay was designed in such a way that integration of pros and cons was one of the last variables that students needed to work on. Long elaborations of females for the variables that were presented to students at the earlier stages such as intuitive opinion and claims in favour and against the topic could have left them with little time to take advantage of the feedback of the learning partners for the integration of the pros and cons, which was almost the last assignment of the students.

Apart from the integration of the pros and cons, female and male students showed no difference with respect to their responses to the argumentative feedback of the learning partners in their revised argumentative essays. The plausible reason that could explain this lack of difference between females and males might be the short duration of the study. The actual peer feedback processes lasted only for less than two hours in the current study. This duration might have been too short to reveal any significant difference between females and males in terms of their responses to the argumentative feedback of learning partners, in their revised written argumentative essays.

Discussions of results for question 3

All students regardless of their gender were able to enhance their knowledge on the content as reflected in their post-test compared with a pre-test. This implies that both females and males benefited from the argumentative peer feedback processes and activities. This is similar to previous studies that emphasize the role of argumentation and peer feedback for fostering students' learning and knowledge on the issue at hand (see Noroozi et al., 2011). Receiving argumentative feedback from the learning partners and also analysing learning partners' arguments leads to a better reflection on the content and understanding of the issue at hand (Bayerlein, 2014; Crisp, 2007). These constructions and receptions of argument followed by peer elaborations, deep cognitive processing, discovering the complementary knowledge of the other students in the group, and clarifications enhanced students' knowledge on the topic (see Noroozi et al., 2013; Schellens, Van Keer, De Wever, & Valcke, 2007). Analysing their group members' essays and providing detailed argumentative feedback on them somewhat raised their awareness of the pros and cons of the topic. Such awareness was then reflected in the post-test assessment of their domain-specific knowledge. Knowledge awareness is an important factor that foster students' domain-specific learning and knowledge construction (see Noroozi et al., 2013; Schreiber & Engelmann, 2010).

The improvement in the content knowledge of the students from pre-test to post-test did not differ between females and males. As explained earlier, participants in this study came from a small power distance society. In such society, both male and female students regardless of their gender typically raise their opinion and discuss conflicts in their own knowledge beliefs (Hsu et al., 2017) which in turn may contribute to their content knowledge understanding on the issue at stake (see Noroozi et al., 2018). The lack of differences between females and males in this study might be related to the equal distribution of the power distance.

Conclusions, limitations, and suggestions for future research

This study explored the extent to which gender factor plays a role for argumentative feedback, essay writing, and learning in higher education. Whilst most of the previous studies focused on gender and argumentation in relation to the students' argumentative discourse activities (see Noroozi et al., 2012 for an overview), this study specifically focused on the role of gender for argumentative peer feedback, essay writing and content learning. This study contributes to accumulating evidence that females and males engage in argumentative peer feedback differently although their essay writing and content learning are not related to their gender. Furthermore, this study concludes that engaging in argumentative peer feedback results in the improvement of all students' essay writing and content learning, regardless of their gender. This study was exploratory in nature and conducted in a real educational setting with high ecological validity and practical relevance. This authentic setting thus has put some limitations and constraints that require further research recommendations.

The literature suggests that what constitutes an effective argumentation might be culturally driven and thus argument patterns in each specific culture may influence written argumentation (Uysal, 2008). This study was conducted in a small power distance society with its own dynamics and that is why in some cases, we could not find large differences between female and male students as we expected based on the literature. It would be interesting to conduct such study with participants from different cultural backgrounds to see how differently or similarly female and male students write argumentative essays, engage in peer feedback processes, and respond to argumentative feedback of their learning partners in large and small power distance societies.

In this study, we adjusted an already developed coding scheme of Noroozi et al. (2016) to measure various aspects of the argumentative feedback quality and argumentative essays. Although this coding scheme has been reported reliable and valid both in the previous and also in the current study, we argue that the small effect sizes and P-values in this study might be related to the limited levels of this coding scheme. Our coding scheme had only three levels ranging between zero to two for each variable of the feedback and argumentative essay. This range might have been too small to reveal any significant differences between female and males students. We, therefore, suggest expanding this coding scheme to a five-point Likert scale ranging from zero to four to check whether similar results would be achieved.

In this study, we explored the role of gender for individual students in the group regardless of the gender distribution in the group. So, the dynamic nature of the group and also the interactions between individual members in the group with similar and different gender types were not explored as such. A study by Jeong and Davidson-Shivers (2006), showed that gender distribution in the group is an influential factor for the way students engage in or avoid argumentative discourse. For example, females tended to post fewer rebuttals to the opposition statements and challenges of females than males, and males tended to post more rebuttals to the challenges of females. Furthermore, males posted twice as many rebuttals in response to disagreements than females. These findings plea for attention on the gender of students when composing learning groups. It would be insightful to explore how heterogeneous or homogenous learning groups in terms of gender perform during argumentative peer feedback and essay writing. Furthermore, future research should explore how males respond to females and vice versa in heterogeneous learning groups. It would

also be interesting to explore how males respond to males and also how females respond to females in homogeneous learning groups. This might have consequences for the formation of the heterogeneous and/or homogeneous learning groups when engaging in argumentative peer feedback and joint essay writing.

In this study, the total duration for the peer feedback processes was short lasting only for about two hours. As discussed before, this could have played a role for not revealing a significant difference between females and males in terms of improvements in the quality of students' argumentative essays and content learning from pre-test to post-test (see Noroozi et al., 2012). A long-term duration study that is expanded over a couple of days may reveal the actual differences between females and males with regard to the dependent variables used in this study. Therefore, long-term duration studies are suggested to see how females and males perform in the online argumentative peer feedback settings.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Omid Noroozi is a faculty member at the Education and Competence Studies Chair Group, Wageningen University, The Netherlands, Wageningen University; his research interest are Collaborative Learning, Computer Supported Collaborative Learning.

Javad Hatami is a faculty member at the faculty of humanities at Tarbiat Modares University, Tehran; his research interests include Learning and Assessment, Educational Technology and E-Learning.

Arash Bayat is a researcher at the Islamic Azad University of Sanandaj, Iran; his research interests concern E-Learning, Online Learning Environments, Virtual and Game-Based Learning.

Stan Van Ginkel is a researcher and lecturer at the Archimedes Institute, University of Applied Sciences, the Netherlands; his research interests concern Educational Psychology, Peer Feedback, Virtual and Augmented Reality.

Harm J. A. Biemans is an Associate Professor at the Education and Competence Studies Chair Group, Wageningen University, The Netherlands; his research interests concern Educational Psychology, Educational Development, and CSCL.

Martin Mulder is retired full Professor at the Education and Competence Studies Chair Group, Wageningen University, The Netherlands; his research interests concern Competence-Based education, Educational Development, and Educational Sciences.

ORCID

Omid Noroozi http://orcid.org/0000-0002-0622-289X
Harm J. A. Biemans http://orcid.org/0000-0003-2955-8211
Martin Mulder http://orcid.org/0000-0002-8729-2477

References

Andrews, R. (1995). *Teaching and learning argument*. London, NY: Cassell.

Asterhan, C. S. C. (2018). Exploring enablers and inhibitors of productive peer argumentation: The role of individual achievement goals and of gender. *Contemporary Educational Psychology*, *54*(1), 66–78.

Asterhan, C. S. C., Schwarz, B. B., & Gil, J. (2012). Small-group, computer-mediated argumentation in middle-school classrooms: The effects of gender and different types of online teacher guidance. *British Journal of Educational Psychology*, *82*(3), 375–397.

Bayerlein, L. (2014). Students' feedback preferences: How do students react to timely and automatically generated assessment feedback? *Assessment & Evaluation in Higher Education*, *39*(8), 916–931.

Caspi, A., Chajut, E., & Sapporta, K. (2008). Participation in class and in online discussions: Gender differences. *Computers & Education*, *50*(3), 718–724.

Coffin, C., & O'Halloran, K. (2008). Researching argumentation in educational contexts: New directions, new methods. *International Journal of Research and Method in Education*, *31*(3), 219–227.

Crisp, B. R. (2007). Is it worth the effort? How feedback influences students' subsequent submission of assessable work. *Assessment & Evaluation in Higher Education*, *32*(5), 571–581.
De Lisi, R., & Golbeck, S. L. (1999). Implications of Piagetian theory for peer learning. In A. M. O'Donnell, & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 3–37). Mahwah, NJ: Lawrence Erlbaum Associates.
De Nisi, A. S., & Kluger, A. N. (2000). Feedback effectiveness: Can 360-degree appraisals be improved? *Academy of Management Executive*, *14*(1), 129–139.
Erkens, G., & Janssen, J. (2008). Automatic coding of dialogue acts in collaboration protocols. *International Journal of Computer-Supported Collaborative Learning*, *3*(4), 447–470.
Gabelica, C., Van den Bossche, P., Segers, M., & & Gijselaers, W. (2012). Feedback, a powerful level in teams: A review. *Educational Research Review*, *7*(2), 123–144.
Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*(3), 445–476.
Hofstede, G. (1993). Cultural constraints in management theories. *Academy of Management Executive*, *7*(1), 81–94.
Hsu, P.-S., Van Dyke, M., & Smith, T. J. (2017). The effect of varied gender grouping on argumentation skills among middle school students in different cultures. *Middle Grades Review*, *3*(2), 1–22.
Jeong, A. C., & Davidson-Shivers, G. (2006). The effects of gender interaction patterns on student participation in computer-supported collaborative argumentation. *Educational Technology Research and Development*, *54*(6), 543–568.
Kellogg, R. T., & Whiteford, A. P. (2009). Training advanced writing skills: The case for deliberate practice. *Educational Psychologist*, *44*(4), 250–266.
Li, Q. (2002). Gender and computer-mediated communication: An exploration of elementary students' mathematics and science learning. *Journal of Computers in Mathematics and Science Teaching*, *21*(4), 341–359.
Lin, S., Liu, E., & Yuan, S. (2001). Web-based peer assessment: Feedback for students with various thinking styles. *Journal of Computer Assisted Learning*, *17*(4), 420–432.
Lizzio, A., & Wilson, K. (2008). Feedback on assessment: Students' perception of quality and effectiveness. *Assessment & Evaluation in Higher Education*, *33*(3), 263–275.
McAlister, S., Ravenscroftw, A., & Scanlon, E. (2004). Combining interaction and context design to support collaborative argumentation using a tool for synchronous CMC. *Journal of Computer Assisted Learning*, *20*(3), 194–204.
Mostert, M., & Snowball, J. D. (2013). Where angels fear to tread: Online peer-assessment in a large first-year class. *Assessment & Evaluation in Higher Education*, *38*(6), 674–686.
Munneke, L., Andriessen, J., Kanselaar, G., & Kirschner, P. (2007). Supporting interactive argumentation: Influence of representational tools on discussing a wicked problem. *Computers in Human Behaviour*, *23*(3), 1072–1088.
Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, *37*(4), 375–401.
Noroozi, O. (2018). Considering students' epistemic beliefs to facilitate their argumentative discourse and attitudinal change with a digital dialogue game. *Innovations in Education and Teaching International*, *55*(3), 357–365. doi:10.1080/14703297.2016.1208112
Noroozi, O., Biemans, H. J. A., Busstra, M. C., Mulder, M., & Chizari, M. (2011). Differences in learning processes between successful and less successful students in computer-supported collaborative learning in the field of human nutrition and health. *Computers in Human Behaviour*, *27*(1), 309–318.
Noroozi, O., Biemans, H. J. A., & Mulder, M. (2016). Relations between scripted online peer feedback processes and quality of written argumentative essay. *Internet and Higher Education*, *31*(1), 20–31.
Noroozi, O., Kirschner, P., Biemans, H. J. A., & Mulder, M. (2018). Promoting argumentation competence: Extending from first- to second-order scaffolding through adaptive fading. *Educational Psychology Review*, *30*(1), 153–176.
Noroozi, O., & Mulder, M. (2017). Design and evaluation of a digital module with guided peer feedback for student learning biotechnology and molecular life sciences, attitudinal change, and satisfaction. *Biochemistry and Molecular Biology Education*, *45*(1), 31–39.
Noroozi, O., Teasley, S. D., Biemans, H. J. A., Weinberger, A., & Mulder, M. (2013). Facilitating learning in multidisciplinary groups with transactive CSCL scripts. *International Journal of Computer-Supported Collaborative Learning*, *8*(2), 289–223. doi:10.1007/s11412-012-9162-z
Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2012). Argumentation-based computer supported collaborative learning (ABCSCL). A systematic review and synthesis of fifteen years of research. *Educational Research Review*, *7*(2), 79–106. doi:10.1016/j.edurev.2011.11.006
Pessoa, S., Mitchell, T. D., & Miller, R. T. (2017). Emergent arguments: A functional approach to analyzing student challenges with the argument genre. *Journal of Second Language Writing*, *38*(1), 42–55.
Prinsen, F. R., Volman, M. L. L., & Terwel, J. (2007). The influence of learner characteristics on degree and type of participation in a CSCL environment. *British Journal of Educational Technology*, *38*(6), 1037–1055.
Prinsen, F. R., Volman, M. L. L., Terwel, J., & Van den Eeden, P. (2009). Effects on participation of an experimental CSCL-programme to support elaboration: Do all students benefit? *Computers and Education*, *52*(1), 113–125.
Qin, J., & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System*, *38*(3), 1–13.

Schellens, T., Van Keer, H., De Wever, B., & Valcke, M. (2007). Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups? *International Journal of Computer-Supported Collaborative Learning*, *2*(2-3), 225–246.

Schreiber, M., & Engelmann, T. (2010). Knowledge and information awareness for initiating transactive memory system processes of computer-supported collaborating ad hoc groups. *Computers in Human Behaviour*, *26*(6), 1701–1709.

Selfe, C. L., & Meyer, P. R. (1991). Testing claims for on-line conferences. *Written Communication*, *8*(2), 163–192.

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189.

Stern, L. A., & Solomon, A. (2006). Effective faculty feedback: The road less traveled. *Assessing Writing*, *11*(1), 22–41.

Sullivan, F. R., Kapur, M., Madden, S., & Shipe, S. (2015). Exploring the role of "gendered" discourse styles in online science discussions. *International Journal of Science Education*, *37*(3), 484–504.

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

Underwood, J., Underwood, G., & Wood, D. (2000). When does gender matter? Interactions during computer-based problem solving. *Learning and Instruction*, *10*(5), 447–462.

Uysal, H. H. (2008). *The interplay between culture and writing: Rhetorical and process patterns in L1 and L2 argumentative writing within a cultural context*. Saarbruecken: VDM Verlag Aktiengesellschaft.

Uysal, H. H. (2012). Argumentation across L1 and L2 writing: Exploring cultural influences and transfer issues. *International Journal of Applied Linguistics*, *9*(1), 151–182.

Van Amelsvoort, M., Andriessen, J., & Kanselaar, G. (2007). Representational tools in computer-supported collaborative argumentation-based learning: How dyads work with constructed and inspected argumentative diagrams. *The Journal of the Learning Sciences*, *16*(4), 485–521.

Van Ginkel, S., Gulikers, J., Biemans, H. J. A., & Mulder, M. (2015). Towards a set of design principles for developing oral presentation competence: A synthesis of research in higher education. *Educational Research Review*, *14*(1), 62–80.

Wingate, U. (2012). Using academic literacies and genre-based models for academic writing instruction. *Journal of English for Academic Purposes*, *11*(2), 26–37.

Wood, N. V. (2001). *Perspectives on argument*. Upper Saddle River, NJ: Prentice Hall.

Wu, S. M. (2006). Creating a contrastive rhetorical stance: Investigating the strategy of problematization in students' argumentation. *Regional Language Centre Journal*, *37*(3), 329–353.

Yang, Y. F. (2010). Students' reflection on online self-correction and peer review to improve writing. *Computers & Education*, *55*(3), 1202–1210.

Yang, Y. F. (2011). A reciprocal peer review system to support college students' writing. *British Journal of Educational Technology*, *42*(4), 687–700.