Effects of academic achievement and group composition on the quality of student-generated questions and online procedural prompt usage patterns

Fu-Yun Yu and Wen-Wen Cheng[*]

*Correspondence:
fuyun@mail.ncku.edu.tw
Institute of Education,
National Cheng Kung
University, No. 1, University
Road, Tainan City 70101,
Taiwan

Abstract

This study aims to examine if and how academic achievement and gender group composition affect the quality of online SGQ and the use patterns of procedural prompts provided to support SGQ activities. Forty-one university sophomores enrolled in an English as a foreign language class participated in a four-week study. All questions generated were categorized based on the revised Bloom's taxonomy for quality evaluation, and a content analysis along the set of integrated online procedural prompts was conducted to reveal usage patterns. Five key findings were obtained: First, the provision of the online procedural prompts served as an efficacious learning scaffold to help the participants at both high- and low-academic achievement levels generate the most questions at high-cognitive levels. Second, based on the results of the Fisher's exact test, no significant relationships were found between academic achievement and the quality of SGQ. Third, the participants in the all-male and mixed-gender groups generated the majority of their questions at high-cognitive levels, whereas the all-female group generated an equal number of questions at both low- and high-cognitive levels. Fourth, no significant relationships between the gender group composition and the quality of SGQ were found according to the chi-square test of independence. Fifth, the results of the content analysis revealed that while some same usage patterns related to online procedural prompts were exhibited by students at both low- and high-academic achievement levels and with different gender group compositions, slightly different usage patterns were observed.

Keywords: Academic achievement, Cooperative learning, Gender group composition, Online procedural prompts, Quality of work, Student-generated questions, Usage patterns

Introduction

Pedagogical values and empirical evidence of student-generated questions (SGQ)

Most traditional classrooms are teacher-centered, and students mainly act as the receivers of instructional content. With the increasingly embraced concepts of student-centered classrooms and learner-as-producer (Arruabarrena et al., 2019), the

student-generated questions (SGQ) approach has been welcomed by educators and validated by empirical studies as a generative learning activity (Yu & Wu, 2020). Research has generally supported the positive effects of SGQ for improving academic achievement (Ellwood & Abrams, 2018; Hardy et al., 2014; Khansir & Dashti, 2014; Poot et al., 2017; Sanchez-Elez et al., 2014; Yu et al., 2015; Yu, & Chen, 2014), comprehension (Hardy, et al., 2014; Song, 2016), engagement, motivation (Davis, 2013; Ellwood & Abrams, 2018; Johnson, 2018; Tho et al., 2020; Yu et al., 2015; Yu & Chen, 2014), and higher-order thinking (Hsu, & Wang, 2018; Idek, 2016; Rooney, 2012; Yu & Chen, 2014). For instance, Johnson's research (2018) found that student-generated review questions activities enhanced the recall, understanding, and learning motivation of undergraduates in English or English Education majors toward course content. Furthermore, student-generated assessment items were considered as high task value, and a significantly positive relationship was found between student engagement in online question-generation and answering and achievement (Poot et al., 2017). An empirical study by Hsu and Wang (2018) on online puzzle-based game learning further demonstrated that the group with an added SGQ component outperformed the game mechanics-only group in enhancing algorithmic thinking skills, engagement experiences, and willingness to participate. Considering the supportive role of SGQ in learning, this study aims to adopt SGQ to facilitate English learning and explore factors, including individual differences in academic achievement and gender group composition when conducting SGQ in cooperative learning that might affect the process and performance of SGQ.

Instructional support with procedural prompts for SGQ

As noted, the learning benefits of the SGQ approach have been confirmed by numerous empirical studies in recent decades. Nonetheless, students have been found to have less experience or confidence in SGQ and consider SGQ to be challenging (Yu, 2009). To help ease the situation and in view of scaffolding theory, the provision of pedagogical support via procedural prompts for SGQ activities has been suggested. Originated from Vygotsky (1978), scaffolding theory accentuates the idea of providing support at different levels to suit the cognitive development of the learner. According to scaffolding theory (Vygotsky, 1978), students can gradually move beyond their current ability level under the guidance of more knowledgeable others or with adequately designed instructional support of different types from external resources.

Significant effects of scaffolding on improving the quality of online learning (Doo et al., 2020), fostering metacognition during problem-solving (Jafarigohar & Mortazavi, 2017), and improving academic performance (Zheng, 2016) have generally been reported. For instance, Doo et al.'s meta-analysis (2020) including studies with 64 effect sizes from 2010 to 2019 showed that computer-based scaffolding had statistically significant effects on cognitive, metacognitive, and affective learning outcomes in higher education. The results of another meta-analysis by Zheng (2016) further indicated that multiple self-regulated learning scaffolds in computer-based learning environments are more effective than one specific scaffold in terms of improving academic performance. Additionally, the results from Jafarigohar and Mortazavi's research (2017) found significant improvements in both individual and socially shared metacognition among 240

English as a foreign language (EFL) learners provided with a combination of structuring and problematizing scaffolding mechanisms in a writing task.

In the context and in support of SGQ, a couple of studies investigated the effects of scaffolding in the form of different procedural prompts on the quality of generated questions (e.g., King, 2002; Yu et al, 2013) and its relationship to learning (e.g., Yu & Pan, 2014). For instance, research conducted by Gelmini-Hornsby et al. (2011) and Yu et al. (2013) both confirmed the supportive effects of generic question stems originally proposed by King (1990, 1992) for promoting better learning. Yu & Pan's study (2014) reported that eighth-graders supported with the online 'the answer is' procedural prompt had better academic and SGQ performance than the no-support group when learning civics and citizenship. In addition to the empirical studies mentioned above, the research, conducted by the authors (Yu & Cheng, 2019) investigating the effects of different procedural prompts on online SGQ performance also found that different procedural prompts had significant effects on the cognitive level dispersion of student-generated question, and most questions generated with the "signal words plus the answer is" integrated procedural prompts fell at the high-cognitive level, while the questions generated with "question stem" procedural prompt fell at the low-cognitive level. In light of the effects of scaffolding procedural prompts in supporting SGQ on learning, the authors would further explore how factors, including individual differences in academic achievement and gender group composition in cooperative learning might affect SGQ performance and online procedural prompts usage patterns.

Individual differences in academic achievement and the quality of SGQ

In pursuit of equalitarian education, individual differences in academic achievement are suggested to be a worthwhile factor to address during the implementation of SGQ (Yu et al., 2005). This is a compelling issue, especially taking into consideration that academic achievement has been widely known to affect learning processes and outcomes (Kaya, 2015). From literature review, most of the previous studies discussed the effects of SGQ on academic achievement (e.g., Ellwood & Abrams, 2018; Hardy et al., 2014; Khansir & Dashti, 2014; Poot et al., 2017; Sanchez-Elez et al., 2014; Yu, & Chen, 2014) or the effects of SGQ with procedural prompts on academic or SGQ performance (e.g., Gelmini-Hornsby et. al., 2011; King, 2002; Yu et al, 2013). However, there has been little discussion about the effects of academic achievement on SGQ (Gorjian et al, 2011; Siegler & Pyke, 2013) and use of online procedural prompts. To tap into both the outcome and process aspects of SGQ with the support of online procedural prompts, the research question on if and in what ways academic achievement affects the quality of SGQ and students' use of online procedural prompts during SGQ learning activities serves as the first focus of this study.

Social support for SGQ via cooperative learning

In addition to providing technological support in the form of various online procedural prompts for SGQ (Yu, 2009), another form frequently adopted in classrooms is social support, specifically, cooperative learning. Cooperative learning highlights the process and benefits of active knowledge construction achieved by allowing students to work in groups where they help one another achieve learning objectives (Johnson

& Johnson, 2009). Studies on cooperative learning have established overwhelming evidence on its effectiveness to enhance learning in a wide array of disciplines. For instance, in the domain of language learning, its effects on promoting reading skills (Khan & Ahmad, 2014; Meng, 2010; Pan & Wu, 2013), speaking skills (Meng, 2010), writing skills (Mahmoud, 2014), and motivation (Marashi & Khatami, 2017; Pan & Wu, 2013) have been proven. Its positive effects on academic achievement have also been found in mathematics (Turgut & Gülşen, 2018), psychology (Tran, 2014) and physical education (Fernández-Espínola et al., 2020), among others. Furthermore, according to meta-analysis studies, cooperative learning has significant effects in terms of improving academic achievement performance (Bertucci et al., 2010; Gillies, 2016; Lou et al., 2001; Turgut & Gülşen, 2018) and in transfer of learning (Pai et al., 2014).

In light of the predominately supportive effects of cooperative learning, recently, researchers have started exploring its use in SGQ and have found confirmatory evidence on the positive relation between cooperative learning and student performance of SGQ activities. For example, in a study by Han and Choi (2018), the cooperative SGQ group was found to achieve a higher score on the comprehension posttest as compared to the individual SGQ group. Another study conducted by Wu et al. (2018) demonstrated that a web-based collaborative SGQ workspace better engaged students in cooperative SGQ activities and better enhanced the quality of generated questions as compared to an individual group arrangement.

Gender group composition in cooperative learning

With these encouraging results from studies on cooperative SGQ, the important role of gender group composition in cooperative learning (Cen, et al., 2016; Harskamp et al., 2008; Mobark, 2014; Takeda & Homberg, 2014; Zhan et al, 2015) must be understood to support the arrangement of pedagogically sound cooperative SGQ. This research topic is worth investigation, especially in light of its current unsettled state. For instance, Cen et al.'s study (2016) found that cooperative learning among heterogeneous gender groups benefited students more than homogeneous groups in terms of learning performance. A study by Harskamp et al. (2008) on solution-seeking behavior revealed that female students in mixed-gender and all-female groups didn't learn to solve physics problems and spent more time asking questions as compared to their male classmates. However, research by Zhan et al. (2015) did not find performance differences for females in either same-gender or mixed-gender groups in computer-based collaborative learning, whereas male undergraduate students were found to perform better in a mixed-gender group. Research conducted by Mobark (2014) showed no significant differences in the academic performance of female and male graduate students in cooperative learning settings for a pretest, posttest, and delayed posttest.

As revealed in the previous discussion, currently, the effects of gender composition on learning are inclusive. However, to the best of the authors' knowledge, no empirical studies are available to shed light on this important issue. Hence, the research question as to if and how gender affects the process and outcomes of SGQ in cooperative learning settings serves as the second focus of this study.

The research questions posed in this study

The learning effects of SGQ with procedural prompts and cooperative learning for the provision of instructional and social support, respectively, are well recognized. Nonetheless, factors further affecting the effects of SGQ, specifically, individual differences in academic achievement and gender compositions in cooperative learning situations are few and far between. Hence, this study was aimed toward examining the respective effects of academic achievement and gender group composition on the quality of SGQ and usage patterns for integrated procedural prompts. The authors expected that individual differences in academic achievement and gender group composition have their respective relationships with the quality of SGQ and online procedural prompts usage patterns.

Specifically, the following four research questions (RQ) are proposed:

RQ#1: Does academic achievement have any relationship with the quality of SGQ?

RQ#2: Does academic achievement have any relationship with online procedural prompts usage patterns?

RQ#3: Does gender group composition have any relationship with the quality of SGQ?

RQ#4: Does gender group composition have any relationship with online procedural prompts usage patterns?

Methods

Participants

Forty-one sophomores (22 males, 19 females) enrolled in a 2-credit hour compulsory EFL class from the College of Management at a National University in southern Taiwan participated in this study. The participants' English competency was classified at the intermediate-level based on the campus-wide standardized Test of English for International Communication (TOEIC) mock test held by the language center at the university. Based on the Common European Framework of Reference for Languages (CEFR) and TOEIC scores, the participants' academic achievement level was classified into three levels: below 350 points was the low level, 351 to 550 points was the medium level, and above 551 was the high level. However, to abide by the chi-square calculation rule while considering approximately equal numbers in the different groups, the students' academic achievement levels were re-grouped to two levels. Those below 450 points were categorized as the low-achieving level, and those above 451 points were categorized as the high-achieving level. Therefore, the participants in the 1st SGQ activity were classified as high-achieving level and low-achieving level.

For the cooperative learning in the 2nd SGQ activity, the participants were allowed to choose two or three of other group members freely for questions and answers discussion. Observing the groups the students formed, there were three types: all-male, all-female and mixed-gender groups.

Instructional design and implementation procedures used in the study

SGQ activities

SGQ activities were integrated in two instructional sessions during a four-week study period after mid-term exam. 'The answer is' with 'signal words' and 'generic question stems' were selected as the procedural prompts for the SGQ activities. The 'signal words' (i.e., who, where, when, and how) procedural prompt was chosen because it is one of the most frequently used and easily learned types of prompts for promoting students' understanding of learning materials (Rosenshine et al., 1996). 'The answer is' procedural prompt proposed by Stoyanova and Ellerton (1996), on the other hand, was targeted in light of its facilitating effects on enhancing academic achievement and SGQ performance (Yu & Pan, 2014) and its relevance to vocabulary acquisition (Yu & Yang, 2014). Finally, a set of question stems proposed by King (1990, 1995) were adapted due to their known effects on promoting performance and their pertinence to the current instructional goal, that is, supporting elaborated responses, as highlighted by King (1990, 1992), to make them attainable through the use of generic question stems.
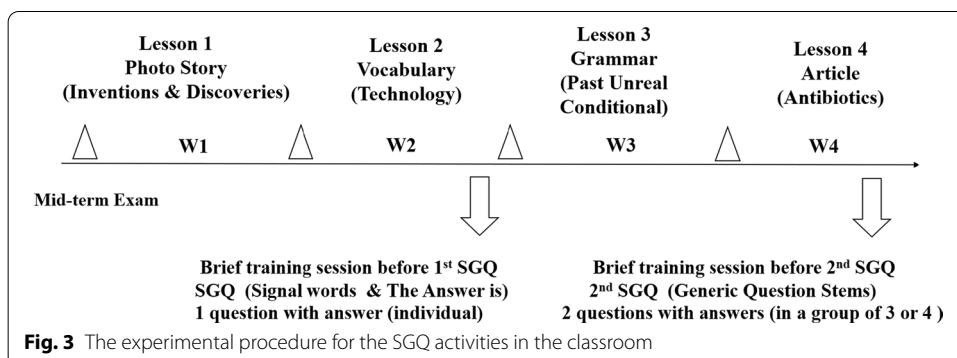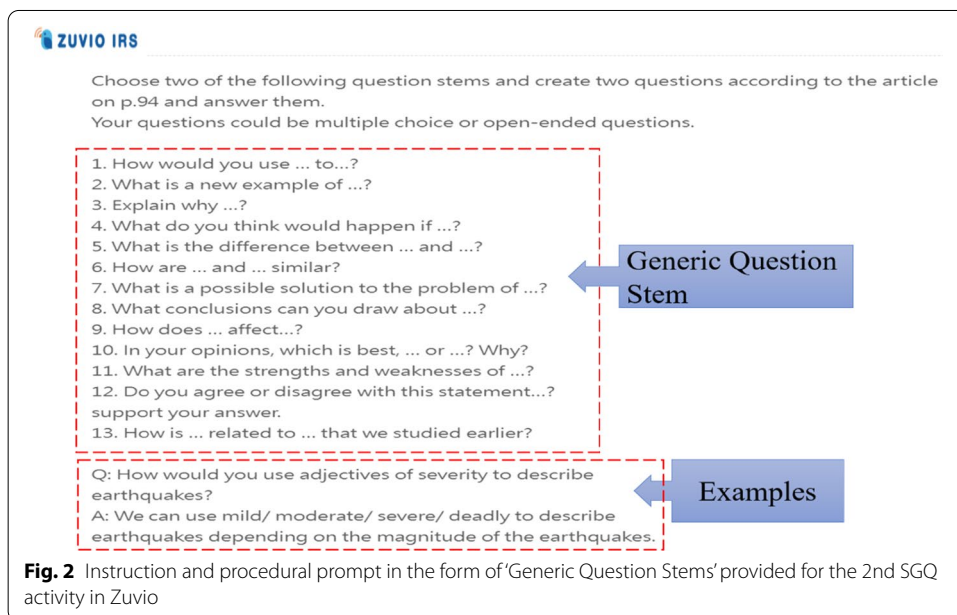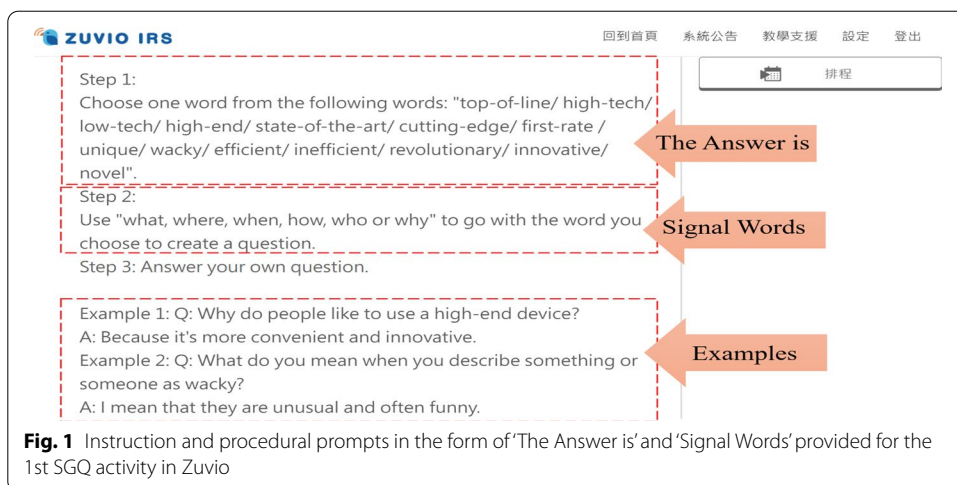
Online learning system supporting SGQ

An online instant interactive system, Zuvio, was adopted for the introduced SGQ activities in class. The participants could access Zuvio using any portable device of their choice (e.g., smartphones, laptop, tablets, etc.) for question/answer generation and submission related to the learned content.

Learning material used in SGQ activities and implementation procedures of SGQ

One unit with four lessons on the topic of *Inventions and Discoveries* from *Top Notch 3* leveled by Pearson publishing as an intermediate-level textbook was selected as the instructional material for the study. The four lessons focused on a photo story (i.e., the topic), vocabulary (on technology), grammar (on past unreal conditional), and an article (on antibiotics), respectively. After lessons 1 and 2, a brief training session on SGQ via Zuvio was scheduled to ensure that the participants were equipped with the relevant knowledge and skills for meaningful engagement in the SGQ activity. Topics introduced for the 1st SGQ activity included how to post questions and answers in Zuvio, how the 'signal words plus the answer is' integrated procedural prompts can be used as a scaffolding device for SGQ (see Fig. 1), and how to generate quality questions. To see the effects of individual differences in academic achievement on SGQ performance and use patterns of procedural prompts, in the 1st SGQ activity, each student generated one question with answer individually corresponding to the instruction delivered on vocabulary in class.

After lessons 3 and 4, the 2nd SGQ activity was scheduled to enhance reading comprehension, and the 'question stems' procedural prompt was introduced. Before the participants engaged in the 2nd SGQ, a brief training session on using the 'question stems' procedural prompt for SGQ (see Fig. 2) was scheduled. The experimental procedure of the SGQ activities is summarized in Fig. 3.

Fig. 1 Instruction and procedural prompts in the form of 'The Answer is' and 'Signal Words' provided for the 1st SGQ activity in Zuvio

Fig. 2 Instruction and procedural prompt in the form of 'Generic Question Stems' provided for the 2nd SGQ activity in Zuvio

Fig. 3 The experimental procedure for the SGQ activities in the classroom

Criteria for classifying the cognitive level of SGQ

To ensure inter-rater reliability, percent of agreement was adopted. The first rater evaluated each SGQ according to the six-level criteria based on the revised Bloom's taxonomy (see Table 1) and rated the SGQ as one of the cognitive levels, and the second rater did the same task. Then, the results of the two raters were compared and any disagreement on the rating was examined and discussed. Finally, the total number of agreement on the rating was divided by the total number of questions rated to reach the percentages of agreement. The results evidenced adequate reliability: 82.96% and 84.38% for the 1st and 2nd SGQ activities, respectively.

Data analysis

For both RQ# 1 and 2 (dealing with the academic achievement factor), as mentioned above, based on CEFR and TOEIC scores, the participants' academic achievement level was classified into two levels: the low-achieving level and the high-achieving level.

The assessment of the quality of SGQ (involved in RQ# 1 and 3) was done by two raters. One rater is a senior university lecturer, and the other is an experienced English teacher. Both assessors independently categorized each of the 123 questions generated by the participants along the revised Bloom's six-cognitive level taxonomy.

Then, for the data analysis for RQ# 1 and 3, the Pearson's chi-square test was applied to analyze whether academic achievement and gender group composition, respectively, has a significant relationship with the quality of SGQ. Given that 33.33% of the cells in the contingency table had a number less than 5, the cognitive levels were grouped into a low level (by combining the bottom three cognitive levels: remember, understand, and apply) and a high level (by combining the top three cognitive levels: analyze, evaluate, and create) to ensure valid chi-square tests and to comply with the calculation rule (i.e., requiring at least 80% of the cells to have an expected count greater than 5).

Finally, for RQ# 2 and 4 (which were concerned with the usage patterns among the integrated online procedural prompts), a content analysis was adopted. The questions generated by the low- and high-achieving participating students during the 1st SGQ activity were analyzed by tallying the number and percentage of use of various 'signal words' (for RQ# 2). Alternatively, the questions generated by the different gender groups

Table 1 Criteria for classifying the cognitive level of SGQ

Dimensions	Definition
Remember	Generating complete questions with answers by recalling learned knowledge or concepts in the textbook
Understand	Generating complete questions with answers describing learned knowledge or concepts in the textbook
Apply	Generating complete questions with answers by using learned knowledge in new examples or situations
Analyze	Generating complete questions with answers by identifying causes or analyzing a problem
Evaluate	Generating complete questions with answers by making judgement about information
Create	Generating complete questions with answers by synthesizing multiple information in new patterns or providing new solution

Table 2 The cognitive levels of SGQ by the low- and high-achieving students ($n = 41$)

		SGQ cognitive levels	
		Low	High
Low-achieving	*f* (%)	4 (16.0%)	21(84.0%)
High-achieving	*f* (%)	2 (12.5%)	14 (87.5%)
Total	*f* (%)	6 (14.6)	35 (85.4%)

Table 3 Use of the 'Signal Words' procedural prompt by the low- and high-achieving students

		Signal words					
		What	Where	When	Why	Who	How
Low-achieving	*f* (%)	7 (27.0%)	0	0	14 (53.8%)	5 (19.2%)	0
High-achieving	*f* (%)	3 (20.0%)	2 (13.3%)	0	9 (60.0%)	0	1 (6.7%)
Total	*f* (%)	10 (24.4%)	2 (4.9%)	0	23 (56.1%)	5 (12.2%)	1 (2.4%)

during the 2nd SGQ activity were analyzed by calculating the frequency and percentage of use of different 'generic question stems' (for RQ# 4).

Results

RQ#1: Relationship between academic achievement and the quality of SGQ

As shown in Table 2, more than 80% of the questions generated by both the low- and high-achieving students were at the high-cognitive level. Taking into consideration that there were two observed values less than 5 in the 2×2 contingency table with small sample sizes, Fisher's exact test was adopted instead of the originally planned Pearson's chi-square test. The results showed that there were no significant relationships between the participants' academic achievement and the quality of SGQ, $p = 1.000 > 0.05$.

RQ#2: Relationship between academic achievement and the usage patterns of online procedural prompts

As shown in Table 3, both 'what' and 'why' signal words were used for SGQ by both the low- and high-achieving students, with 'why' being used most frequently (by more than half of the respective groups), followed by 'what.' Furthermore, it was noted that 'when' was never used by either group. Despite the fact that two same patterns were used by the high- and low-achieving students, some different usage patterns were observed. Explicitly, 'who' was used exclusively by the low-achieving group, whereas 'how' and 'where' were used only by the high-achieving group.

RQ#3: Relationship between the gender group composition and the quality of SGQ

As shown in Table 4, more of the generated questions fell into the high-cognitive level in the case of the all-male and mixed-gender groups while an equal distribution of the generated questions was at the low- and high-cognitive levels in the case of the all-female group. The results of the chi-square test of independence indicated no significant relationships between gender group composition and cognitive levels of SGQ, $p = 0.443 > 0.05$.

Table 4 The cognitive levels of SGQ by gender group composition (*n* = 82)

		SGQ cognitive levels	
		Low	**High**
All-male	*f* (%)	14 (38.9%)	22 (61.1%)
All-female	*f* (%)	12 (50.0%)	12 (50.0%)
Mixed-gender	*f* (%)	7 (31.8%)	15 (68.2%)
Total	*f* (%)	33 (40.2%)	49 (59.8%)

Table 5 Use of 'Generic Question Stems' by different group compositions

Question Stems	All-male *f* [rank]	All-female *f* [rank]	Mixed-gender *f* [rank]	Total *f* (%) [*rank*]
1. How would you use … to…?	2		4	6 (7.3%)
2. What is a new example of …?				
3. Explain why …?	3	6 [1]	9 [1]	18 (22.0%) [2]
4. What do you think would happen if …?	4	6 [1]	7 [2]	17 (20.7%) [3]
5. What is the difference between … and …?	14 [1]	4	2	20 (24.4%) [1]
6. How are … and … similar?				
7. What is a possible solution to the problem of …?				
8. What conclusions can you draw about …?				
9. How does … affect…?	2	5		7 (8.5%)
10. In your opinions, which is best, … or …? Why?				
11. What are the strengths and weaknesses of …?	7 [2]	3		10 (12.2%)
12. Do you agree or disagree with this statement…? support your answer	4			4 (4.9%)
13. How is … related to … that we studied earlier?				
Total	36	24	22	82 (100%)

RQ#4: Relationship between gender group composition and the usage patterns of online procedural prompts

As shown in Table 5, among the 13 question stems, as a whole, three question stems (including no. 3, 4, and 5) were used by all three different gender composition groups while six of the stems (including no. 2, 6, 7, 8, 10, and 13) were used by none of the gender groups. Furthermore, among the seven used question stems, the three used most frequently by all three groups were no. 5, 3, and 4, in that order.

As for differences in the usage pattern among the three gender group compositions, the all-male group used more question stems (a total of seven stems) than the all-female group (five) and the mixed-gender group (four). In addition, question stem no.5 was used most by the all-male group whereas question stem no.3 was used most by both the all-female and mixed-gender groups. Finally, question stem no.12 was used only by the all-male group.

Discussion and conclusions

SGQ has been promoted due to its positive learning effects on cognitive, affective, and social development (Yu & Wu, 2020). Various instructional arrangements either in the form of procedural prompts (Gelmini-Hornsby et. al. 2011; King, 2002; Yu, 2009; Yu

et al, 2013; Yu & Pan, 2014) or social support through cooperative learning (Han & Choi, 2018; Wu et al., 2018) have been suggested and attested to further promote its learning efficacy. In light of the possible effects of individual differences in academic achievement (Gorjian et al., 2011; Kaya, 2015; Siegler & Pyke, 2013) and gender group composition in a cooperative learning situation (Cen, et al., 2016; Harskamp et al., 2008; Mobark, 2014; Takeda & Homberg, 2014; Zhan et al, 2015), issues regarding if and how such factors may affect SGQ were examined in this study. Specifically, individual differences in academic achievement and gender group composition were targeted, and their respective relationships to the quality of SGQ and usage patterns of integrated online procedural prompts were examined in this study to determine the outcome and process aspects associated with SGQ. The authors expected that individual differences in academic achievement and gender group composition have their respective relationships with the quality of SGQ and online procedural prompts usage patterns. The following section will discuss and explain the findings of this study.

First, for RQ#1, in terms of the outcome aspect of SGQ, the results of the Fisher's exact test did not substantiate significant relationships between the participants' academic achievement and the quality of SGQ. In other words, the obtained results did not corroborate the findings of Kaya (2015) on effects of individual differences, where high-achievers generated more higher-order questions than low-achievers. In this study, both high- and low-achievers generated a predominate and similar percentage of questions at a higher cognitive level. This finding is somewhat surprising. However, a comparison of the implementation procedures used in these two studies directed the authors to one possible reason for this difference. It may be the provision of the procedural prompts in this study intended to help direct the participants' attention related to generating questions toward the high-cognitive level and thus eliminated the gap between students at different English achievement levels. Specifically, being provided with an explicit set of 'signal words' with 'the answer is' procedural prompts coupled with concrete examples (see Fig. 1) helped guide the participants at both high- and low-achievement levels to generate questions with answers that demanded analyzing, evaluating, and creating on the basis of the learning content (rather than merely remembering, understanding, and applying what they had acquired from the content). This finding reflects and is in alignment with the results of Bergey (2014), Yu et al. (2013), and Yu and Pan (2014) in that, procedural prompts acting as a scaffolding strategy for SGQ is effective for promoting learning. That is, although no significant relationships between the participants' academic achievement and the quality of SGQ were found, the supportive function of SGQ is confirmed.

Furthermore, for RQ#3, this study did not confirm a significant relationship between gender group composition and the quality of SGQ. The obtained results differed from the findings of Harskamp et al. (2008) and Zhan et al. (2015), who found that students performed differently in different group compositions. As noted and explained in the previous paragraph, the authors conjectured that this may have been due to the additional support provided in the form of generic question stems in terms of alleviating possible effects that may arise from different gender compositions.

In terms of the process aspect of SGQ in RQ#2 and RQ#4, some same usage patterns of online procedural prompts were exhibited by students at both low- and high-academic

achievement levels and those in different gender groups. At the same time, some different usage patterns were observed in terms of the total number, degree of exclusivity, and preferred use of some procedural prompts. For instance, in terms of the total number of procedural prompts used, the high-achieving students used more types of signal words than the low-achieving students, and the all-male group used more question stems than the other two gender composition groups. In terms of the exclusive use of procedural prompts, 'how' and 'where' were referred to only by the high-achieving group, and the 'Do you agree or disagree with this statement…? support your answer' question stem was used only by the all-male group. For the preferred use of procedural prompts, the 'What is the difference between … and …?' and 'What are the strengths and weaknesses of …?' question stems were used most frequently by the all-male group, whereas the 'Explain why …?' and 'What do you think would happen if …?' question stems were used most by both the all-female and mixed-gender groups. This phenomenon may be understood and reflect the findings of Bromley (2013), Charoento (2016), O'Malley et al. (1990) and Taheri et al. (2020), who conducted research on the selection and use of learning strategies by students at different achievement levels throughout the language learning process. As demonstrated in O'Malley et al.'s (1990) study, when completing a language task, learners with higher language proficiency adopted diverse learning strategies to reach their goals. Similarly, Taheri et al.'s (2020) research findings indicated that high- and low-achievers applied different language learning strategies, with high achievers employing more strategies than low-achievers. With that said, it should be noted that in this study, despite the fact that slightly different usage patterns of SGQ procedural prompts were observed, this did not lead to significantly different quality of SGQ (in terms of cognitive level dispersion) between individuals at different academic performance levels, or among groups of different gender compositions.

The significance of this study

This study provided preliminary empirical data on the relationships between academic performance and gender group composition on the quality of SGQ and usage patterns of online procedural prompts. What is more, this paper highlighted the potential educative effects of the provision of online procedural prompts as efficacious scaffolds for students at different academic achievement levels and those in different gender group compositions in terms of their ability to generate questions at higher-cognitive levels.

Suggestions for instructors

The current study confirmed that integrating an explicit set of online procedural prompts (either in the form of 'signal words' with 'the answer is,' or 'generic question stems') to support SGQ activities is effective in terms of directing students to generate questions at higher-cognitive levels for language learning, regardless of individual differences in academic achievement and gender group composition under cooperative learning situations. On the basis of the findings obtained in this study and other studies supporting the use of procedural prompts (e.g., Gelmini-Hornsby et. al. 2011; King, 2002; Yu et al, 2013; Yu & Pan, 2014), it is suggested that instructors give an explicit set of procedural prompts during SGQ activities to achieve high-quality question-generation.

Limitations of this study and suggestions for future studies

This study was limited by a small sample size and the short duration of the study. Studies with a larger sample size and extended study periods in the future are needed for external generalizability purposes. Moreover, as noted by King (2002), the quality of questions asked by students were controlled by the generic question stems provided, which in turn influenced their cognitive level. While this study revealed that some different usage patterns were exhibited by the participants, questions as to why different online procedural prompts are considered and used by students at different academic achievement levels and in different gender group compositions during SGQ activities are yet to be answered. Hence, this requires further investigation. This line of investigation would be better if conducted using a qualitative research method (e.g., in-depth interviews) to gain insight that will help instructors set up distinct guidelines on the provision of SGQ procedural prompts for students with individual differences or cooperative groups with different gender compositions.

Declarations

Competing interests
The authors declare that they have no competing interests.

References
Arruabarrena, R., Sánchez, A., Blanco, J. M., Vadillo, J. A., & Usandizaga, I. (2019). Integration of good practices of active methodologies with the reuse of student-generated content. *Journal of Educational Technology in Higher Education, 16*, 10. https://doi.org/10.1186/s41239-019-0140-7
Bergey, B. W. (2014). *Student-generated questions during chemistry lectures: Patterns, self-appraisals, and relations with motivational beliefs and achievement*. [Unpublished Dissertation at Temple University].
Bertucci, A., Conte, S., Johnson, D., & Johnson, R. (2010). The impact of size of cooperative group on achievement, social support, and self-esteem. *Journal of General Psychology, 137*(3), 256–272.
Bromley, P. (2013). Active learning strategies for diverse learning styles: Simulations are only one method. *Political Science & Politics, 46*(4), 818–822.
Cen, L., Ruta, D., Powell, L., Hirsch, B., & Ng, J. (2016). Quantitative approach to collaborative learning: Performance prediction, individual assessment, and group composition. *International Journal of Computer-Supported Collaborative Learning, 11*(2), 187–225.
Charoento, M. (2016). Individual learner differences and language learning strategies. *Journal of Contemporary Educational Psychology, 7*, 57–72.
Davis, T. A. (2013). Connecting students to content: Student-generated questions. *Bioscene: Journal of College Biology Teaching, 39*(2), 32–34.
Doo, M. Y., Bonk, C., & Heo, H. (2020). A meta-analysis of scaffolding effects in online learning in higher education. *International Review of Research in Open and Distributed Learning, 21*(3), 60–80.
Ellwood, R., & Abrams, E. (2018). Student's social interaction in inquiry-based science education: How experiences of flow can increase motivation and achievement. *Cultural Studies of Science Education, 13*(2), 395–427.

Fernández-Espínola, C., Abad Robles, M. T., Collado-Mateo, D., Almagro, B. J., Castillo Viera, E., & Giménez Fuentes-Guerra, F. J. (2020). Effects of cooperative-learning interventions on physical education students' intrinsic motivation: A systematic review and meta-analysis. *International Journal of Environmental Research and Public Health*, *17*(12). https://doi.org/10.3390/ijerph17124451

Gelmini-Hornsby, G., Ainsworth, S., & O'Malley, C. (2011). Guided reciprocal questioning to support children's collaborative storytelling. *International Journal of Computer-Supported Collaborative Learning, 6*(4), 577–600.

Gillies, R. M. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education*, *41*(3), 3, article 3. https://ro.ecu.edu.au/ajte/vol41/iss3/3/

Gorjian, B., Moosavinia, S. R., Ebrahimi Kavari, K., Asgari, P., & Hydarei, A. (2011). The impact of asynchronous computer-assisted language learning approaches on English as a foreign language high and low achievers' vocabulary retention and recall. *Computer Assisted Language Learning, 24*(5), 383–391.

Han, S. Y., & Choi, Y. H. (2018). Post-reading question-generation activities and cooperative learning in EFL reading. *English Teaching, 73*(2), 37–58.

Hardy, J., Bates, S. P., Casey, M. M., Galloway, K. W., Galloway, R. K., Kay, A. E., Kirsop, P., & McQueen, H. A. (2014). Student-generated content: Enhancing learning through sharing multiple-choice questions. *International Journal of Science Education, 36*(13), 2180–2194.

Harskamp, E., Ding, N., & Suhre, C. (2008). Group composition and its effect on female and male problem-solving in science education. *Educational Research, 50*(4), 307–318.

Hsu, C. C., & Wang, T. I. (2018). Applying game mechanics and student-generated questions to an online puzzle-based game learning system to promote algorithmic thinking skills. *Computers & Education, 121*, 73–88.

Idek, M. S. (2016). Q-matrix as a method in promoting student-generated questions to develop critical thinking skills. *Journal on Technical and Vocational Education, 1*(1), 30–38.

Jafarigohar, M., & Mortazavi, M. (2017). The impact of scaffolding mechanisms on EFL learners' individual and socially shared metacognition in writing. *Reading & Writing Quarterly, 33*(3), 211–225.

Johnson, C. N. (2018). What did we learn last class? An exploratory study of student-generated review questions. *PASAA: Journal of Language Teaching and Learning in Thailand*, *56*, 65–79.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*(5), 365–379. https://doi.org/10.3102/0013189X09339057

Kaya, S. (2015). The effect of the type of achievement grouping on students' question generation in science. *The Australian Educational Researcher, 42*(4), 429–441.

Khan, S. A., & Ahmad, R. N. (2014). Evaluation of the effectiveness of cooperative learning method versus traditional learning method on the reading comprehension of the students. *Journal of Research & Reflections in Education, 8*(1), 55–64.

Khansir, A. A., & Dashti, J. G. (2014). The effect of question-generation strategy on Iranian EFL learners' reading comprehension development. *English Language Teaching, 7*(4), 38.

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal, 27*(4), 664–687.

King, A. (1992). Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist, 27*(1), 111–126.

King, A. (1995). Guided peer questioning: A cooperative learning approach to critical thinking. *Cooperative Learning and College Teaching, 5*(2), 15–19.

King, A. (2002). Structuring peer interaction to promote high-level cognitive processing. *Theory into Practice, 41*(1), 33–39.

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71*(3), 449–521.

Mahmoud, M. M. A. (2014). The effectiveness of using the cooperative language learning approach to enhance EFL writing skills among Saudi university students. *Journal of Language Teaching and Research, 5*(3), 616.

Marashi, H., & Khatami, H. (2017). Using cooperative learning to boost creativity and motivation in language learning. *Journal of Language and Translation, 7*(1), 43–58.

Meng, J. (2010). Cooperative learning method in the practice of English reading and speaking. *Journal of Language Teaching and Research, 1*(5), 701.

Mobark, W. M. (2014). Effect of using cooperative learning strategy on graduate students' academic performance and gender difference. *Journal of Education and Practice, 5*(11), 64–70.

O'malley, J. M., O'Malley, M. J., Chamot, A. U., & O'Malley, J. M. (1990). *Learning strategies in second language acquisition*. Cambridge University Press.

Pai, H., Sears, D., & Maeda, Y. (2014). Effects of small-group learning on transfer: A meta-analysis. *Educational Psychology Review, 27*, 79–102.

Pan, C. Y., & Wu, H. Y. (2013). The cooperative learning effects on English reading comprehension and learning motivation of EFL freshmen. *English Language Teaching, 6*(5), 13–27.

Poot, R., De Kleijn, R. A., Van Rijen, H. V., & van Tartwijk, J. (2017). Students generate items for an online formative assessment: Is it motivating? *Medical Teacher, 39*(3), 315–320.

Rooney, C. (2012). How am I using inquiry-based learning to improve my practice and to encourage higher order thinking among my students of mathematics? *Educational Journal of Living Theories, 5*(2), 99–127.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66*(2), 181–221.

Sanchez-Elez, M., Pardines, I., Garcia, P., Miñana, G., Roman, S., Sanchez, M., & Risco, J. L. (2014). Enhancing students' learning process through self-generated tests. *Journal of Science Education and Technology, 23*(1), 15–25.

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology, 49*(10), 1994.

Song, D. (2016). Student-generated questioning and quality questions: A literature review. *Research Journal of Educational Studies and Review, 2*, 58–70.

Stoyanova, E., & Ellerton, N. F. (1996). A framework for research into students' problem posing. In P. Clarkson (Ed.), *Technology in mathematics education* (pp. 518–525). Melbourne: Mathematics Education Research Group of Australasia.

Takeda, S., & Homberg, F. (2014). The effects of gender on group work process and achievement: An analysis through self- and peer-assessment. *British Educational Research Journal, 40*(2), 373–396.

Taheri, H., Sadighi, F., Bagheri, M. S., & Bavali, M. (2020). Investigating the relationship between Iranian EFL learners' use of language learning strategies and foreign language skills achievement. *Cogent Arts & Humanities, 7*(1), 1710944.

Tho, P. D., Lai, C. H., & Navratil, O. (2020). Effect of behavior patterns of accessing learning materials on learning performance in student-generated question activities. *Journal of Internet Technology, 21*(1), 135–146.

Tran, V. D. (2014). The effects of cooperative learning on the academic achievement and knowledge retention. *International Journal of Higher Education, 3*(2), 131–140.

Turgut, S., & Gülşen Turgut, İ. (2018). The effects of cooperative learning on mathematics achievement in Turkey: A meta-analysis study. *International Journal of Instruction, 11*(3), 663–680.

Vygotsky, L. S. (1978). Socio-cultural theory. *Mind in Society, 6*, 52–58.

Wu, C. P., Chen, Y. C., & Wu, S. L. (2018, January). Development of a collaborative learning space for student generated-question strategy. In M. Krishnamurthi and M. Iinuma (Eds), *Proceedings of the 6th international conference on information and education technology* (pp. 102–105), Osaka, Japan.

Yu, F. Y. (2009). Scaffolding student-generated questions: Design and development of a customizable online learning system. *Computers in Human Behavior, 25*(5), 1129–1138.

Yu, F. Y., Chang, Y. L., & Wu, H. L. (2015). The effects of an online student question-generation strategy on elementary school student English learning. *Research and Practice in Technology Enhanced Learning, 10*(1), 1–16.

Yu, F. Y., & Chen, Y. J. (2014). Effects of student-generated questions as the source of online drill-and-practice activities on learning. *British Journal of Educational Technology, 45*(2), 316–329.

Yu, F. Y., & Cheng, W. W. (2019). The effects of different procedural prompts on online student-generated question performance in terms of cognitive levels. In *27th International conference on computers in education, ICCE 2019* (pp. 399–404). Asia-Pacific Society for Computers in Education.

Yu, F. Y., Liu, Y. H., & Chan, T. W. (2005). A Web-based learning system for question-posing and peer assessment. *Innovations in Education and Teaching International, 42*(4), 337–348.

Yu, F. Y., & Pan, K.-J. (2014). Effects of student question-generation with online prompts on learning. *Educational Technology and Society, 17*(3), 267–279.

Yu, F. Y., Tsai, H. C., & Wu, H.-L. (2013). Effects of online procedural scaffolds and the timing of scaffolding provision on elementary Taiwanese students' question-generation in a science class. *Australasian Journal of Educational Technology, 29*(3), 416–433.

Yu, F. Y. & Wu, W. S. (2020). Effects of student-generated feedback corresponding to answers to online student-generated questions on learning: What, why, and how? *Computers & Education, 145*, 103723.

Yu, F. Y., & Yang, Y. T. (2014). To see or not to see: Effects of online access to peer-generated questions on performance. *Educational Technology and Society, 17*(3), 27–39.

Zhan, Z., Fong, P. S., Mei, H., & Liang, T. (2015). Effects of gender grouping on students' group performance, individual achievements and attitudes in computer-supported collaborative learning. *Computers in Human Behavior, 48*, 587–596.

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review, 17*(2), 187–202.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.