A Comparative Study of Student Question Generation in Web-based and Paper-based Learning Environment*

Kyung-Mi O (Dongduk Women’s University)


Over the last few decades, researchers have been investigating the uses of student-generated questions (SGQs) in multifarious areas. In understanding the potential benefits of SGQ, this action research is designed to examine a group of students’ perceptions after experiencing interventions of web-based student generated questions (e-SGQs) and paper-based student-generated questions (PB-SGQs) in the classroom. For this research, ten students were recruited to construct test items for their own mid-term preparation. The student generated items were split fairly evenly with about half of the items being e-SGQs and the other half being PB-SGQs. The two modes were compared through the analyses of a survey, a group interview, and individual in-depth interviews. The data analyses reveal that the participants perceived their experiences of generating questions to be very helpful for understanding their class content. Although many of them clearly recognized the advantages of e-SGQs as an educational tool, most of the participants chose PB-SGQ over e-SGQ. This result can be explained by the survey and interview data which illustrate that the respondents preferred the test taking strategies that they were used to such as marking, underlining, and skimming through the test items on a paper.

**Key words**: alternative assessment, student generated questions, online assessment, student-centered assessment, paper-based assessment, test mode

doi: 10.15702/mall.2019.22.3.102

I. INTRODUCTION

In the current educational practice which emphasizes students’ pro-active learning, it is important to encourage students to be highly engaged in the learning process as well as the assessment process (Papinczak, et al., 2012). The importance of the student participation in the assessment has also been highlighted in Stefani’s study (1998), which introduced the idea of the

* This work was supported by the research fund of Dongduk Women’s University.
partnership between students and teachers for empowering the learners in the assessment process. In this respect, having students construct their own assessment items is one such endeavor.

This approach of asking students to design their own questions is a classroom assessment technique (CAT) which is often called student-generated test questions (SGQs) (Angelo & Cross, 1993). The strengths of using SGQs in educational settings have been addressed by a number of researchers (Hawe & Dixon, 2014; Kafrani & Afshari, 2017; Lan & Lin, 2011). Despite the potential benefits of the SGQs, however, there has been a dearth of research studies investigating how the students perceive the SGQs. Moreover, there have been even fewer studies examining the relationship between students’ attitudes and the SGQs with different test modes.

Due to the development of multiple test delivery media, various assessments have been formed using different modes including paper, computer, and mobile devices. The SGQs also are expected to be delivered through diverse modes. Since the testing mode is believed to influence test-takers’ motivation and testing performance (Wenemark, Persson, Brage, Svensson, & Kristenson, 2011), and assessments are known to largely affect students’ attitudes toward their learning (Biggs, 2003), it seems important to find out how the students perceive the experience of different assessment modes.

Given the situations, as action research, this study is motivated to investigate a group of students’ perceptions after experiencing the interventions of web-based student generated question (e-SGQ) and paper-based student-generated question (PB-SGQ). To address the purpose, answers to the following four research questions were sought:

1) How do the students perceive their experiences of generating test items?
2) How do the students perceive taking the e-SGQ?
3) How do the students perceive taking the PB-SGQ?
4) Which mode of testing do the students prefer for the SGQ items?

II. LITERATURE REVIEW

1. Student Question Generation

For the last four decades, student-question generation (SQG) has been examined as a way to increase students’ learning outcomes (Cohen, 1983; King, 1992). The idea of SQG is rooted
on generative learning theory (Wittrock, 1974; 1990) and elaborative interrogation (Pressley, McDaniel, Turnure, Wood, & Ahmad, 1987). According to the generative learning theory, learners combine new ideas into their memory by associating new information with what they had previously learned in order to promote their learning experience. The elaboration techniques involve linking new information with familiar topics of relevance, highlighting the distinct differences among similar ideas, and building an overview in diverse forms (Hoffman, 1997; Reigeluth, 1999). Both the generative learning theory and the elaboration technique emphasize the role of learners’ activity of making new information meaningful by associating the new information with what they had previously learned, and this aspect has been accepted by many researchers.

Given that, special attention has been paid to self-generated elaboration (Wittrock, 1990). Compared with elaboration provided by external sources (e.g., a teacher, a textbook, etc.), self-generated elaboration has been reported to be more productive in learning (Wittrock, 1990). According to King (1992), learners usually find the activities of engaging in their own elaboration easier than processing external elaboration and therefore more effective in its recall. In this respect, asking students to generate questions has been seen to be more useful than other kinds of external elaborations in efficient recall of information. In addition, the advantages of SQG have been reported in various research studies.

SQG has been addressed to be effective in increasing student motivation (Barak & Rafaeli, 2004; Chin & Brown, 2002; King, 1992; Lan & Lin, 2011; Wilson, 2004) and autonomy (King, 1992; Marbach-Ad & Sokolove, 2000). SQG is also known to be useful in allowing instructors to diagnose their students’ weaknesses (Kafrani & Afshari, 2017), assisting students to find out important points (Nicol, 2010), and to digest specific test contents (Havnes, Smith, Dysthe, & Ludvigsen, 2012), and encouraging students’ high-level thinking (King, 1992; Papinczak, et al., 2012) and/or deep-learning (Hawe & Dixon, 2014; Roscoe & Chi, 2007). With the exception of a few studies (Byun, Lee, & Cerruto, 2014; Kafrani & Afshari, 2017) reporting no significant result, SQG has been widely considered to motivate students to voluntarily participate in discussions (Barak & Rafaeli, 2004; Chin & Brown, 2002; King, 1992; Lan & Lin, 2011; Wilson, 2004) and to assist them to link their prior knowledge to new learning points (Nicol, 2010).

2. Online Assessment vs. Paper-based Assessment

Current trends in language assessments have been moving toward the online mode from the paper-based mode, which is partly related to the advantages that online assessments offer
(Nikou & Economides, 2016). The benefits include ease of administration, immediacy of scoring and reporting, flexibility of test scheduling and ubiquity of test settings.

In understanding the usefulness of the online assessments, many researchers have examined the relationships between the test modes and the students’/teachers’ perceptions (Chen & Jang, 2010; Deutsch, Herrmann, Frese, & Sandholzer, 2012; Huff, 2015; Hwang & Chang, 2011; Jamil, Tarig, & Shami, 2012; Macario, 2009; Macedo-Rouet, Ney, Charles, & Lallich-Boidin, 2009; Sun, 2014). Although there are some exceptions reporting students’ preference toward paper-based mode (Macedo-Rouet et al., 2009) and displaying no significant difference between the two modes (Calik & Simsek, 2010), many studies have displayed students’ positive attitudes toward the assessments with information technology (Deutsch, et al., 2012; Macario, 2009; Romero, Ventura, & De Bra, 2009).

For instance, Jamil, et. al., (2012) have compared teachers’ perceptions about computer-based tests and paper-based tests. For the study, teachers were grouped into seven categories (i.e., gender, department, designations, qualifications, teaching experiences, computer training certificates, and computer-based examination experiences) and were surveyed in regards to their perceptions toward the two modes of tests. Through the study, the researchers concluded that, with minimal exceptional cases, the overall teachers’ attitudes were positive toward computer-based test systems. They also added that teachers equipped with computer training certificates or with some experience of computer-based tests were more positive toward the computer-based tests. The results of the study suggest that some experience of using a computer while teaching may have been a reason for some of the positive attitudes toward computer-based tests.

In another comparative study, Nikou and Economides (2016) examined the effect of three modes of self-assessment—paper-based, computer-based and mobile-based—on students’ motivation and achievement. For the study, the researchers provided each of the three groups of students with assessments in paper-based, computer-based and mobile-based versions. The three types of assessments were made with the same questions that the classroom teacher designed, and the students were tested on their understandings of the class contents before and after the experiment. In addition, the three groups of students responded to a questionnaire specifically designed to measure their motivation before and after the experiment. Through the analysis of data, the researchers found the students responded more positively toward the computer- and mobile-based assessments than the paper-based versions. Based on the analyses, the researchers concluded that computers, as well as mobile modes, could be used in the place of paper-based assessments.
The study of Nikou and Economides is meaningful in that the researchers conducted the comparative analyses investigating the effect of the relatively new modes of assessment on students’ motivation and learning outcomes. Moreover, their research designs and the ways of analyses are judged to be reasonable. However, one aspect that seems uncertain is the reason why the researchers thought the assessments they employed were self-assessment. According to Paris and Paris (2001), self-assessment means a procedure that the students learn to assess what they know, understand what they need to learn, and observe their own learning progress to reach their educational goal. Boud and Falchikov (1989) also defined it as “involvement of learners in making judgements about their own learning” (p. 529). Based on these definitions, the assessments in the three modes employed in Nikou and Economides’s study, all with 10 multiple-choice, true-false, or fill-in-the blank questions designed by the instructor, seem to be closer to a homework assignment rather than self-assessment. Nevertheless, the researchers’ idea of implementing the self-assessment procedure on a computer and mobile mode seems to be very well-timed and suitable in that the digital technology provides multiple functions which are often presumed to increase the students’ motivation and involvement.

3. Student Engagement and Online Assessment

The digital technology provides students with flexibility of time and space and often contains multiple functions for stimulating their motivations and engagement for study. In understanding this aspect, Chen (2010) designed a self- and peer-assessment system called Mobile Assessment Participation System (MAPS) on the platform of Personal Digital Assistants (PDAs). For the experiment, the researcher had 37 students use MAPS for assessing themselves and their peers. Through the study, the researcher claimed that both teachers and students could benefit from MAPS in terms of flexibility and efficiency involving assessment and learning procedures. Despite the positive side of the experience, the students also revealed their concerns involving issues of objectivity and reliability of the assessment. Chen’s assessment system using a PDA seems relatively outdated compared to the current popularity of smartphones. However, the study is worth noting in that it clearly displays the importance of students’ perceptions when it comes to designing and using assessment tools that highlight students’ involvement.

In another study utilizing technology, Yu and Chen (2014) investigated the usefulness of student-generated questions on the web (i.e., online drill-and-practice activities) for increasing the students’ academic performance and motivation. Through a five-week experiment with 145
participants, the researchers compared the academic performance and motivation of the three groups: a) SQG and answering teacher-generated questions, b) online SQG and solving student-generated questions, and c) online SQG alone (comparison group). The researchers did not find online SQG and solving student-generated questions to be more beneficial to academic performance and motivation as compared with online SQG and answering teacher-generated questions or the online SQG. Despite the results, the researchers have claimed that the combination of online SQG and answering student-generated questions does have some educational potential which needs further investigation.

Therefore, research studies using current IT tools for SQG appear to be timely. Motivated by many SQG studies and available technology to aid the activity, this research study was aimed at understanding the students’ perceptions toward their experiences of generating test items and solving the items both on online and paper-based modes.

III. METHOD

1. Action Research Design

For the study, a practical action research design was adopted. This method is usually employed in situations where teacher-researchers intend to improve the practice of education through the systematic investigation of a local issue (Lesha, 2014). Thus, the action research design was selected because the teacher-researcher intended to find if there was an efficient way to increase students’ understandings of the class content by examining how the SGQs were perceived in local educational contexts. Moreover, the use of an action research design has been known to be advantageous in that it improves student achievement through more effective instructions and school administrations (Cohen & Manion, 1980; Elliot, 1991; Kemmis, 1981; Stenhouse, 1975).

Despite the benefits, studies using action research are often limited in that the sample sizes are not large enough for statistical significance, restricting quantitative analyses to descriptive rather than generalizable. In fact, much research data from action research are qualitative as the focus of the research cannot be detached from the local research settings. Due to this lack of external validity, the outcome from the current action research was not expected to have the same results with other populations. In addition, special attention was needed for interpreting the data.
As a way to overcome the limitations, this study took the form of “multilevel research” (Tashakkori & Teddlie, 1998, p. 48) which was basically combining different methods to form a methodological triangulation. Specifically, in this study, three kinds of multiple data collection methods including survey, group interview and individual in-depth interview procedures were employed.

2. Participants

The participants of this study consisted of all ten students who were enrolled in one of the teacher-researcher’s English teaching courses at a women’s university. The participants in the study were all female in their early twenties majoring in English. Since the course is purposely designed for only ten percent of the English major students aiming to acquire teaching certificates along with their bachelor’s degree, only a limited number of students usually take the course. The contents of the course cover understanding curriculum design and teaching methods and constructing assessment tools and teaching materials. Thus, the students were expected to have the minimal level of knowledge in creating test items and judged to be ideal subjects for the study.

In spite of the selection of the suitable participants, using ten individuals for a research study is a very small sample. Moreover, on the day when the survey was to be administered, one student was absent, leaving only nine students to respond to the survey. The same student did not show up for the individual in-depth interview session either, although she participated in the group interview session allowing the researcher to gather the interview data from all ten participants. This small sample size limits the current study significantly as it increases the chances of error skewing the outcomes, and thus again, it should not be generalized in other settings. In sum, for this action research, ten students participated in constructing SGQs and a group interview, and nine of them were involved in the survey and in-depth individual interviews.

3. Procedure

At the onset of the study, the participants learned about the student question generation and diverse test tasks for test design. Throughout the two weeks of group and individual work, they designed their own SGQs for preparing the mid-term test and submitted the test items to the teacher-researcher.
In order to create the SGQs, the ten students in the class were divided into two groups consisting of five members in each. The students in each group were asked to design at least five items from a chapter they chose from the two chapters to be tested on the mid-term. The students made 72 test items with various test tasks including T/F, multiple-choice, multiple response, open-ended (type-in), matching, and fill-in tasks.

The 72 items were reviewed by the researcher to examine if there were overlapping items, any typos, items with incorrect keys or distractors, and any other potential problems. Two overlapping questions were deleted and some items with incorrect keys or problems were corrected by the researcher, which resulted in 70 items. Out of the 70 items, 40 were from Chapter 1 and 30 were from Chapter 2. The 40 items were formed into e-SGQs and the remaining 30 were made into PB-SGQs.

The two types of tests were used by the students multiple times in groups, individually in class, and at home over a four week period. In the 6th week, a questionnaire, a group interview session, and individual in-depth interviews were administered and the gathered data were analyzed.

Both group and individual interview data were collected by the researcher using Korean language, the mother tongue of both researcher and the interviewees. Segments of the data were transcribed and translated for the provision in the following result section.

4. Materials

1) Web-based Student-generated Questions (e-SGQs)

The e-SGQs consisted of 40 student-generated questions from Chapter 1. For creating the test on the web, e-Learning software, iSpring QuizMaker 8 was chosen.

This app was selected because it bears resemblance with the MS office packet, which is familiar to the researcher and the students. Also, the app provides the users with multiple functions including diverse test task choices, autoscoring, feedback, and multiple-trial option.

The created e-SGQs were published on the iSpring Cloud and the Cloud address was posted on the e-class community bulletin. The tests on the Cloud were judged to provide the students with easy access from their personal computers and smartphones anywhere, anytime (see Figure 1).
2) Paper-based Student-generated Questions (PB-SGQs)

The PB-SGQs were composed of thirty student generated test items from Chapter 2. The items were first solved in class and the same items with answers and feedback were also given thereafter (see Figure 2).
3) Questionnaire

A survey was made to measure the students' perceptions toward their experiences of SQG and taking the e-SGQ and PB-SGQ. The instrument was constructed with a 5 point Likert-scale items allowing the researcher to quantify the survey results. It is mainly comprised of three sections: students’ experiences of 1) designing test items, 2) taking the e-SGQ/PB-SGQ, 3) functions of e-SGQ and students’ preference (see Table 1).

The questions were partly inspired and adopted from Gordon (2015), O (2018), and Song (2015). To be more specific, four items (#1~2, #7~8) from Song (2015, p. 160), four (#9, #12~14) from O (2018, p. 97) and two (#18, #22) from Gordon (2015, p. 32) were directly adopted; four items (#3~6) from Song (2015), two (#10~11) from O (2018) were used with some modification; and lastly, six items (#15~17, #19~20) were newly added.

Before the administration, the instrument was reviewed by four individuals: an expert in TESOL with about thirty years of teaching experience in higher education and three students who have experience designing tests. After the review, the final version was administered to the
participants. For the statistical analysis of the quantitative data from the survey, SPSS 25 was used. The Cronbach’s coefficient alpha value calculated was 0.892 indicating an acceptable reliability coefficient compared to the standard acceptability of Cronbach’s coefficient alpha of 0.7 (Nunnally, 1978).

(Table 1) Survey Questions with Framework and Specifications

<table>
<thead>
<tr>
<th>Survey Framework</th>
<th>Item #</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Students’ experiences of SQG</td>
<td></td>
</tr>
<tr>
<td>A. Question generation experience</td>
<td></td>
</tr>
<tr>
<td>I learned how to make good questions.</td>
<td>1</td>
</tr>
<tr>
<td>I am able to make good questions.</td>
<td>2</td>
</tr>
<tr>
<td>B. Perceived Outcome</td>
<td></td>
</tr>
<tr>
<td>Making questions was helpful for me to understand the class content.</td>
<td>3</td>
</tr>
<tr>
<td>Making questions was helpful for me to read the content critically.</td>
<td>4</td>
</tr>
<tr>
<td>Making questions allowed me to think about important concepts in the class content.</td>
<td>5</td>
</tr>
<tr>
<td>C. Satisfaction Degree</td>
<td></td>
</tr>
<tr>
<td>I enjoyed making questions.</td>
<td>6</td>
</tr>
<tr>
<td>I was satisfied with making questions.</td>
<td>7</td>
</tr>
<tr>
<td>I would like to make questions for the final test.</td>
<td>8</td>
</tr>
<tr>
<td>II. Students’ experiences of taking PB-SGQ/e-SGQ</td>
<td></td>
</tr>
<tr>
<td>D. Perceived Outcome</td>
<td></td>
</tr>
<tr>
<td>Solving PB-SGQ/e-SGQ was helpful for me to understand the class content.</td>
<td>9-1, 9-2</td>
</tr>
<tr>
<td>Solving PB-SGQ/e-SGQ was helpful for me to read the content critically.</td>
<td>10-1, 10-2</td>
</tr>
<tr>
<td>Solving PB-SGQ/e-SGQ allowed me to think about important concepts in the class content.</td>
<td>11-1, 11-2</td>
</tr>
<tr>
<td>E. Satisfaction Degree</td>
<td></td>
</tr>
<tr>
<td>I enjoyed solving PB-SGQ/e-SGQ in general.</td>
<td>12-1, 12-2</td>
</tr>
<tr>
<td>I am satisfied with solving PB-SGQ/e-SGQ.</td>
<td>13-1, 13-2</td>
</tr>
<tr>
<td>I would like to solve PB-SGQ/e-SGQ for the final test.</td>
<td>14-1, 14-2</td>
</tr>
<tr>
<td>III. Functions of e-SGQ &amp; Students’ Preference</td>
<td></td>
</tr>
<tr>
<td>F. Functions of e-SGQ</td>
<td></td>
</tr>
<tr>
<td>* _________ function was helpful for my understanding the class content.</td>
<td></td>
</tr>
<tr>
<td>Autoscoring</td>
<td>15</td>
</tr>
<tr>
<td>Multiple trial</td>
<td>16</td>
</tr>
<tr>
<td>Feedback</td>
<td>17</td>
</tr>
<tr>
<td>G. Students’ Preference</td>
<td></td>
</tr>
<tr>
<td>Which test did you prefer? Why?</td>
<td>18</td>
</tr>
<tr>
<td>I would choose ______ for my test items. Why?</td>
<td>19</td>
</tr>
<tr>
<td>For educational purposes considering the test impact, I would choose ______. Why?</td>
<td>20</td>
</tr>
<tr>
<td>I think ______ better serve the goals of my test items. Why?</td>
<td>21</td>
</tr>
<tr>
<td>Do you feel the testing format changed the difficulty of the tests? Why?</td>
<td>22</td>
</tr>
</tbody>
</table>
4) Interview Guide

To understand the students’ perceptions in detail and supplement the questionnaire data, an interview guide with a basic checklist was designed for the semi-structured interviews in both groups and individually. The use of an interview guide with a basic checklist has been known to be beneficial for researchers to have “in-depth probing while permitting the interviewer to keep the interview within the parameters traced out by the aim of the study” (Berg, 2007, p. 39). The interview guide in this study basically included questions about the advantages and disadvantages of each test mode and suggestions for future research and test methods.

IV. RESULTS

1. Question Generation

To the first research question inquiring about the students’ experiences of generating questions, most of the students responded to the survey in a positive way (see Table 2).

<table>
<thead>
<tr>
<th>Survey Item #</th>
<th>Question generation experience</th>
<th>Perceived Outcome</th>
<th>Satisfaction Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>I learned how to make good questions.</td>
<td>4.11 (.60)</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>2. I am able to make good questions.</td>
<td>3.89 (.78)</td>
<td>4.00</td>
</tr>
<tr>
<td>B.</td>
<td>3. Making questions was helpful for me to understand the class content.</td>
<td>4.56 (.53)</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td>4. Making questions was helpful for me to read the content critically.</td>
<td>4.56 (.73)</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td>5. Making questions allowed me to think about important concepts in the class content.</td>
<td>4.89 (.33)</td>
<td>5.00</td>
</tr>
<tr>
<td>C.</td>
<td>6. I enjoyed making questions.</td>
<td>4.00 (1.00)</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>7. I was satisfied with making questions.</td>
<td>4.00 (1.00)</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>8. I would like to make questions for the final test.</td>
<td>4.00 (.87)</td>
<td>4.00</td>
</tr>
</tbody>
</table>

As seen in Table 2, to the survey questions regarding question generation experience, perceived outcome and satisfaction degree, the students mostly provided positive responses with the means ranging from 3.89 to 4.89 (and the medians\(^1\) of 4.00 or 5.00). The lowest
among the listed means was 3.89 with a standard deviation of .78 indicating somewhat moderately positive attitude.

The qualitative data from the group interviews support the quantitative results. The following excerpts from the comments of Ann, Briana and Cecile display that the students perceived the SQG's outcome positively:

Ann: I felt that I should understand the content thoroughly to design test items. For instance, in order to design open-ended questions, I had to have complete understanding about the content. But at the same time, the items shouldn’t be too easy ‘cause I don’t want them to be giveaway items. I liked the experience overall because I had to study the content clearly in order to make the test items.

Briana: I realized that I had to have thorough understanding about the test content to make test items.

Cecile: Making the questions allowed me to read the content more critically. I think I could read each word or expression more attentively and critically.

The comments from Dana, Elena, and Felicity from in-depth individual interviews also demonstrate the positive aspects of SQG:

Dana: I had to understand about the content in detail to make test items. Particularly, making X questions in the OX test tasks was difficult. Honestly, I haven’t prepared that much for the mid-term before. But since I had this experience, I feel kind of confident about the mid-term.

Elena: While I was designing the SGQs, I realized that there weren’t too many important things to cover in a chapter. I studied at least this one chapter so well for designing the SGQs. So for me, preparing the mid-term was not so burdensome or overwhelming experience because of the (SQG) experience.

Felicity: Since I knew that I reviewed this one chapter thoroughly, I felt less anxious about that chapter before (the mid-term).

The three students’ responses illustrate that the experience of SQG was helpful for them for understanding the content, and it allowed them to lower the test anxiety to some degree. Therefore, the three kinds of data all seem to display that the students perceived the experience of SQG positively.

---

1) In considering the ordinal data and the small sample size with non-parametric distribution, it is appropriate to use medians rather than means. However, as seen in much research reporting the ordinal data analyses with the mean, this study presents the descriptive data with the mean data for convenience.

2) All the names used in the interview excerpts are pseudonym to maintain the participant’s anonymity.
2. The e-SGQ versus the PB-SGQ

To the second and third research questions asking about the students’ perception toward taking the e-SGQ/PB-SGQ, the students showed positive attitudes toward both e-SGQ and PB-SGQ (see Table 3).

<table>
<thead>
<tr>
<th>Survey Item #</th>
<th>D. Perceived Outcome</th>
<th>E. Satisfaction Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>9-1, 2</td>
<td>4.44 (.53)</td>
<td>4.22 (.67)</td>
</tr>
<tr>
<td>10-1, 2</td>
<td>3.89 (.60)</td>
<td>4.56 (.53)</td>
</tr>
<tr>
<td>11-1, 2</td>
<td>4.00 (.50)</td>
<td>4.44 (.53)</td>
</tr>
<tr>
<td>12-1, 2</td>
<td>4.22 (.67)</td>
<td>4.22 (.67)</td>
</tr>
<tr>
<td>13-1, 2</td>
<td>4.33 (.71)</td>
<td>4.33 (.71)</td>
</tr>
<tr>
<td>14-1, 2</td>
<td>3.78 (.83)</td>
<td>4.56 (.73)</td>
</tr>
</tbody>
</table>

As shown in Table 3, to the survey questions asking about the perceived outcome and satisfaction level of taking e-SGQ, the students mostly provided positive responses with the means ranging from 3.89 to 4.44. To the counterpart questions for the PB-SGQs, the participants also rendered favorable results with the lowest mean of 4.22 to the highest mean of 4.56. Overall, the students seem to have been satisfied with their experiences of the two test modes although the means of PB-SGQs were slightly higher. These quantitative results were confirmed by the comments from the in-depth individual interviews:

*Gina:* I enjoyed solving both e-SGQs and PB-SGQs. I enjoyed them because I felt I was the owner of this learning process.

*Heidi:* Usually when I prepare for my mid-term or final tests, I often miss some important points from my studying, but solving other students’ questions was helpful for me to catch those missed points.

The interview data seem to demonstrate that the students positively perceived the experience of taking both e-SGQs and PB-SGQs, but many students including Irene expressed their preference of PB-SGQs over e-SGQ.

*Irene:* I think both modes were good. But I personally preferred PB-SGQs to e-SGQ because I could solve the items
This type of students’ preference to the PB-SGQs was further investigated and confirmed through the survey items #18 to 21 and the interview data in the following students’ preference section.

3. Students’ Preference

As to the last research question inquiring about the students’ preference, except the item asking about serving the educational purposes, most of the respondents chose PB-SGQ over e-SGQ (see Table 4).

<table>
<thead>
<tr>
<th>(Table 4) Students’ Preference</th>
<th>e-SGQ (%)</th>
<th>PB-SGQ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. Students’ Preference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Which test did you prefer?</td>
<td>33.3</td>
<td>66.7</td>
</tr>
<tr>
<td>19. I would choose ______ for my test items.</td>
<td>22.2</td>
<td>77.8</td>
</tr>
<tr>
<td>20. For educational purposes considering the test impact, I would choose ______.</td>
<td>77.8</td>
<td>22.2</td>
</tr>
<tr>
<td>21. I think ______ better serve the goals of my test items.</td>
<td>33.3</td>
<td>66.7</td>
</tr>
</tbody>
</table>

Despite similar levels of positive attitudes toward the e-SGQ shown in Table 3, as seen in Table 4, the majority of the students (66.7%) responded that they prefer PB-SGQ to e-SGQ (#18). Furthermore, even more students (77.8%) answered that they would choose PB-SGQ over e-SGQ (#19).

The students in the individual interviews explained the reason by saying that they wanted to read on the paper rather than on a screen on their smartphone or computer. Regarding this, Ann, Brianna, and Cindy had the following to say:

Ann: PB-SGQs are easier for my eyes to see. Because I’m used to paper-based tests, I prefer it over the e-SGQs.

Brianna: I want to see all the test items on one page to see the whole picture of the test.

Cindy: I like marking and underlining on paper while solving the test.

The comments from Ann, Brianna, and Cindy suggest that the students would choose PB-SGQ over e-SGQ because they wanted to work on a paper-based mode, which they were
used to. In the group interview, the majority of the students also displayed their preference for the paper assessment over the online format. Dana from the group interview and Elena from the individual interview shared their views:

_Dana: Somehow, (as a test-taker), I think I can concentrate better with a paper-based version. So I think the test takers would be able to figure out my intention of designing the test items._

_Elena: It was easier for me to read the contents on paper because I could reread them multiple times (as many times as I want) with better concentration…._

Both Dana and Elena believed that they could concentrate better with PB-SGQs. The relationship between test takers’ level of concentration and the test mode needs further investigation. And yet, the students’ comments seem to share the same vein with Myrberg and Wiberg (2015) highlighting the importance of students’ habits and attitudes of learning medium in their learning processes.

To item #21 asking about serving the goals of their own test items, many respondents answered that they prefer PB-SGQ to e-SGQ. In the group interview, Gina and Heidi reported the followings:

_Gina: Since PB-SGQ allows the test takers to concentrate and read better than e-SGQ, the goals of my test items would be expressed better in the paper version._

_Heidi: Because I intended to create multiple test items that are related, the paper-based version would allow me to express my goals more clearly. The web-based version seems to present the test items in somewhat random and/or choppy way._

Gina’s comment illustrates that the paper-based version would better serve the goals of her test designs since she believes it is easier to concentrate on PB-SGQs. Heidi preferred PB-SGQs because she wanted a test format which displayed multiple items at once. The comments from Gina and Heidi can provide the educators and test designers with important implications and insights for future research investigating the designs of new online interface.

To the last item, #22 (see Table 1) asking if the testing format changed the difficulty of the tests, most of the respondents answered affirmatively with 77.8 percent of the students providing positive responses. Many of them explained that the feedback and multiple-trial functions of e-SGQ helped them feel the test tasks were easier than test tasks from the paper-based version. Only 22.2 percent of the students replied that they found no difference in the difficulty of the test caused by the mode of the test. This result illustrates the potential benefit of e-SGQs in allowing the students to perceive test items less difficult, and follow-up
research should be carried out to further investigate the aspect.

To question #20, the majority of the students (77.8%) responded that they perceived e-SGQ was more helpful to serve an educational purpose. Although PB-SGQs also came along with the feedback and answers, and the test designers themselves explained about the items to the class after solving the items, many of them preferred the feedback function of e-SGQs. They reported that the feedback of e-SGQs was extremely helpful as it presented detailed explanations for each test item every time they chose wrong answers. Due to this immediate feedback function, the students seemed to firmly believe the positive aspect of e-SGQ for increasing educational efficiency. Irene and Jessica in the individual interviews made the following remarks:

Irene: *Due to the immediate feedback function that the e-SGQ provided, the educational function such as correcting the misinterpretations of the class content and supplementing the incomplete knowledge of it seems to have been better served in e-SGQ.*

Jessica: *Because e-SGQ allowed the test takers to solve the test multiple times, it led me to think about the test items repetitively and learn them better. I think I could remember the content longer (through e-SGQ) because of the repetition.*

The students’ group and interview data were further confirmed by the data collected from the survey questions of #15, 16 and 17 (see Table 5).

<table>
<thead>
<tr>
<th>F. <em>Function of e-SGQ</em></th>
<th>Mean (SD)</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>_________ function was helpful for my understanding the class content.</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Autoscoring</td>
<td>4.33 (.71)</td>
<td>4.00</td>
</tr>
<tr>
<td>16. Multiple trial</td>
<td>4.78 (.44)</td>
<td>5.00</td>
</tr>
<tr>
<td>17 Feedback</td>
<td>4.78 (.44)</td>
<td>5.00</td>
</tr>
</tbody>
</table>

To the questions about the usefulness of the functions provided in e-SGQ, the students provided positive responses largely. To the questions inquiring about the usefulness of each function of e-SGQ (autoscoring, multiple trial and feedback), the means of the respondents were 4.33, 4.78, and 4.78 all suggesting favorable views. These positive responses seem to have been closely related to the students’ positive remarks in the usefulness of e-SGQ as an educational instrument. In sum, however, to the last research question asking about the students’ preference for designing SGQs, majority of the students reported that they would
choose PB-SGQ over e-SGQ.

V. CONCLUSION

This study has examined a group of students’ attitudes toward the experience of generating their own questions and then answering the questions both online and offline through e-SGQs and PB-SGQs. The results indicated that the students perceived their experiences of SQG positively for increasing content knowledge. Additionally, the study revealed that the students positively perceived solving the questions through the two modes of e-SGQ and PB-SGQ. Although many of them plainly expressed the benefits of e-SGQ in serving the educational goals, the students still preferred PB-SGQ over e-SGQ. This was because the paper-based version provided the students with many benefits of itself. It appeared that many of the students still wanted to solve the items on paper by marking, underlining, and skimming through multiple items.

In spite of the findings, this study is limited in many aspects. Since it took a practical action research design using merely ten participants, the data analyses had to be limited to descriptive statistics without a further significance test, and it had to be restricted to the interpretation within the given context without further generalization to other settings. Moreover, the experiment was conducted for a relatively short period of time, which lasted only seven weeks. Notwithstanding the weaknesses, by employing the method triangulation, this study was designed to overcome the limitations in researching in a real educational setting.

Through the analyses, this action research confirms Nicole (2010), and Havnes et al. (2012) in that, at least to the participating students’ views, SQG was seen to increase the content knowledge. Although the current data displayed the students’ positive views toward both online and offline modes of SGQs, the study illustrated that the majority of the students preferred PB-SGQs to e-SGQs confirming Macedo-Rouet et al. (2009). Despite the students’ inclination toward PB-SGQs, the students distinctively expressed their satisfaction about the efficiency of e-SGQs in its serving an educational goal because of the feedback and multiple trial functions. Thus, this kind of information is expected to be of useful to educators and program developers in the educational settings. With a more systematic and rigorous research design, many more future research investigating SQG studies should be carried out.
REFERENCES


Comparative Study of Student Question Generation in Web-based and Paper-based Learning Environment


Applicable levels: secondary, tertiary education

Author: O, Kyung-Mi (Dongduk Women’s University); kmo@dongduk.ac.kr

Received: July 31, 2019
Reviewed: August 20, 2019
Accepted: September 15, 2019