Master's Thesis in Graduate School of

Library, Information and Media Studies

An Initial Study of Learning Methods to Improve the Quality of Student-Generated Questions

September 2020

201826099

Izhar Almizan Wahono

An Initial Study of Learning Methods to Improve the Quality of Student-Generated Questions

Student No.: 201826099
Name: Izhar Almizan Wahono

Student-generated questioning or question posing activity proved to have a positive impact on students' learning as students are enabled to develop higher-order thinking skills and seek richer knowledge. However, students are having trouble in creating high level questions which is why they might not be able to develop higher-order thinking skills. One way to solve this issue is to implement the activities from the creative processes, that recommends a way to generate a high-level idea by gathering inspiration, generating the idea, and also improving its quality through refining the idea. In this study, we implemented online learning with question posing and question refinement activities to improve students' generated questions quality. Utilizing off-the-shelf online tools from the Internet, we proposed a learning strategy for generating better questions by incorporating the activities in creative processes collaboratively and individually. The activities are *gathering* knowledge, question *generation*, and question *refinement*. These activities are performed respectively.

We examined our proposed learning strategy using a between-subject design with different numbers of students involved (individual, two students, and three students). The students are coming from different academic backgrounds, which are generally not familiar with our selected topic for the learning material.

Based on our initial pilot study we found out that most of the students are able to improve the quality of other students' questions and their own questions in the refinement stage based on questioning rubric by Taboada and Guthrie. Moreover, we also found that students refining questions collaboratively in the pair condition have no significant difference in terms of quality level compared to students refining individually. Whereas when students are refining questions collaboratively in the triad condition, they can improve the questions to a better quality.

In addition, from the questionnaire and interview results, the students claimed that the question generation and refinement activities helped them gain deeper understanding and helped them memorize the learning material.

Academic Advisors: Principal: Tomoo Inoue
Secondary: Sayan Sarcar

An Initial Study of Learning Methods to Improve the Quality of Student-Generated Questions

Izhar Almizan Wahono

Graduate School of Library,
Information and Media Studies

University of Tsukuba

September 2020

Contents

List of Figures

Chapter 1

Introduction

Empirical studies on student-generated questioning (SGQ) have been well-documented and upheld the effects of facilitating students' cognitive and affective development. Various benefits of SGQ were reported in the literature; for instance, SGQ assists students in identifying the crucial aspects of the learning content and focusing on its essential elements. Students can deepen their understanding of the study material, reflect on and understand the information they receive, concentrate on using the learned knowledge and skills in solving new problems and situations, rather than just "acquiring" the knowledge/skills, and finally, they can relate the newly acquired information to their prior knowledge. SGQ is also deemed useful in helping students monitor and appraise their comprehension level, reinforcing their degree of confidence regarding the subject matter, helping students embrace a positive attitude in learning, and growing their interest and skills in the problem-solving activity following the SGQ activity. Through this, students can develop divergent thinking and enhance their learning strategies [1, 2, 3, 4, 5].

Furthermore, student-generated questioning is found to be a practical technique for learning since it supports an active process. Once students actively construct relations between their prior knowledge and the learning material, students build-up their understanding and comprehension. During the process of question generation, students allotted their attention and cognition on generating meaning through finding relevant information within the material and connecting it with their knowledge, which helped enhance their comprehension [6, 7, 8]. Through the student-generated questioning activity, students are engaged in analyzing the content, making relationships with prior knowledge, evaluating it, and constructing personal knowledge.

However, despite the benefits of student-generated questioning, one study found that students have difficulty creating high-level questions [9]. In addition, other researchers reported a low number of high-level questions during implementation [10]. Moreover, studies have reported that using student-generated questioning strategy has appeared to have different results depending on whether it is performed in groups or individually [11, 12, 13]. One way to solve this issue is to implement the activities from the creative processes (see Section 2.8), that recommends a way to generate a high-level idea by gathering inspiration, generating the idea, and also improving its quality through refining the idea by manipulating, reorganizing, structuring, altering, and or removing information [14]. Our study investigates the activity of refining student posed questions individually and collaboratively in regards to increasing the quality level of the questions.

Question refinement has been used to improve the quality of a question in a certain parameter, different researches using different approaches have been conducted in order to investigate the improvement of questions' quality. For instances, King applied question refinement through a continuous process of self-questioning and work in groups to ask the question and answer each other's questions [11]. Moreover, Denny et al.[15] and Yu [16] have also demonstrated this technique by asking students to evaluate multiple-choice questions created by others based on the quality of the question, distractors and the explanation why the answer is correct. This technique was used in one of the newest studies in 2020, [17] in which the students can modify each other's questions and claim ownership of the modified question by adding justification to why the previous question needed to be refined.

Researchers in machine learning have also tried to improve the quality of a question in terms of clarity, spelling, and grammar. For instance, Elgohary [18] constructed a dataset that contains standalone questions to which context was added, in order to improve the clarity, such as "What happened in 1983?" to "What happened to AnnaVissi in 1983?". Furthermore, Liu [19] proposed a unified model called Qrefine to refine ill-formed questions to well-formed questions.

Inspired by both question-posing and refinement techniques, we came up with the following research question:

> **RQ:** How the generated questions could change and evolve through question refinement activity?

To answer our research question, we proposed a learning strategy that enables students to pose questions to a particular lecture video and refine the existing questions to have better quality. The procedure will provide guidance as a form of scaffolding to students for creating questions at a better-quality level.

Chapter 2

Related Work

2.1 Collaborative Student-Centered Learning

Student-centered learning (SCL) is one of the learning approaches where the learning process is centered on a student's inquiry to achieve knowledge with the help of a teacher as a facilitator. The traditional teaching method is a more teacher-centered learning environment: the teachers choose the content, look up the references, restructure the information, and then deliver the knowledge to the students in the classroom [20]. The problem with the teacher-centered approach is that students become passive information receivers from teachers or tutors without any active thinking.

One of the concepts related to the SCL is self-regulated learning (SRL), which has been receiving increased attention in the educational field, especially in higher education. In SRL, students have the ability to control and monitor their learning processes and determine how to locate suitable learning resources [21]. As the MOOCs on the Internet is getting popular recently as a distance learning environment, making SRL more prevalent as MOOCs give everyone freedom of learning time and pace without any physical restriction such as the traditional classroom.

In a collaborative learning context, similarly to SCL and SRL, learners share and communicate knowledge in order to achieve mutual "learning goals," they are actively involved in acquiring knowledge through various practices, such as having discussions, seeking information, and interchanging opinions. The acquired knowledge is shared among learners instead of it being owned by one learner [22]. The collaborative learning process is used in the context of student-generated questions to help students acquire "higher-order thinking skills" as it requires students to use their metacognitive abilities.

2.2 Student Generated Questions

Rosenshine et al. [2] defined the activity of question generation as both a cognitive strategy and a metacognitive strategy. In this sense, the process of asking questions is considered a way to build up learners' comprehension as they focus on main ideas, and to verify the learners' level of understanding in order to confirm if the content is assimilated.

In the context of social-cognition, question-generation is considered a helpful activity and a crucial element of student debate in 'talking science' [23, 24]. Additionally, asking questions can lead students to discuss and establish dialogues from different viewpoints, encourage them to contemplate and examine issues from different perspectives (pros and cons). It can promote argumentation and critical thinking in science which is considered an essential process in which students can acknowledge insufficient or inadequate reasoning and false assumptions, build-up hypotheses, come-up with explanations, determine evidence which either supports or disproves a hypothesis, assesses options logically, and try to link discordant ideas.

Diverse research was conducted on student-generated questions, such as PeerWise [15], a tool that students use to create multiple-choice questions (MCQs) and respond to their peers' questions. This tool can be used to support the learning process in different ways, as students are asked to create questions meant to align with the course's learning outcome. The process of coming up with questions and writing an explanation for the answer to their question can help students improve their understanding of the course content.

Moreover, Hirai et. al. [25] examined how students can improve their understanding on what they learned by learner- centered learning with question-posing. Based on the results, they reported that the students who actively posed questions scored higher score than the students who were passive in the answer test. It also reported that the learners who asked higher quality questions scored higher on the both pre- and post-test than the learners who posed low-quality.

Furthermore, Saarinen et. al. [26] suggested a process called Adaptive Tool-Driven Conception Generation, in which they developed a tool that was implementing this process. The tool can explicitly optimize the process for multiple-choice questions that divide students' opinions. It can create questions similar to expert-designed questions, produce novel questions that can identify potential student misinterpretations, and calculate statistical estimation and probability of misconceptions.

Previous studies on student-generated questions discussed the process of asking questions and suggested different tools to help students generate questions [15, 26], however, the generated questions could be distinguished in terms of its quality level.

2.3 Types and Quality of Questions

In different studies, the following questions were discussed: "What types of questions are considered 'quality questions'?" "Which qualities or properties qualify some questions as better or worse than others?". For instance, Graesser and Person distinguished between shallow, short answer questions addressing and describing the content and high-level questions comprising inferences, multi-step reasoning, applying an idea to new knowledge, synthesizing new ideas from different information resources and evaluating new claims [27].

In another study, Elder and Paul [28] describe the following criteria as helpful to qualify a question as good: "clarity, accuracy, precision, relevance, depth, breadth, and logic". Similarly, using a process called "reciprocal questioning", King reported the successful results he obtained using a process called 'Reciprocal questioning' in which students are instructed to use high-level questions in asking each other [29]. These types of questioning sessions were compared with equal time open-ended discussions, and although the latter's answers were longer, they were approximately all classified "low level". The questioning groups were deemed remarkably better in terms of 'critical thinking' and 'high-level elaboration'.

Fisher [30] declared that the planning and also the use of quality questions are related to 'critical thinking skills', in which he perceives learning critical thinking as involved with questioning. Fisher encourages teaching students' 'generic' questions, such as:

- What is the main idea here? How would you compare this with . . .? But how is that different from . . .? Now, can you give me a different example? How does this affect . . .?

In Browne and Keeley [31], critical thinking is defined in terms of the awareness of and ability to ask 'critical questions', such that critical thinking is the:

1. Awareness of a set of interrelated critical questions,
2. Ability to ask critical questions at appropriate times, and
3. Desire to actively use these critical questions. (p. 2)

In his study, De Jesus [32] differentiated between "confirmation questions" and "Transformation questions", based on which he came up with a taxonomy of questions, the class measures of each type of question is based on constructing meaning and building-up frameworks of understanding. According to the author, confirmation questions are meant to clarify information, to differentiate between facts and opinions, to examine issues, and ask for definitions and examples. These questions help in deciding whether the information is applicable and admissible, also in understanding the basis for its inclusion in a certain context or setting as well as deciding on its worth and place. On the other hand, Transformation questions are used to identify logical steps, to omit, to understand thought's structures, and to defy accepted reasoning. These questions imply a reorganization of the student's understanding. The student is assumed to go deeper within the ideas, to be hypothecio-deductive, to examine the overlapping of knowledge domains, and to go beyond what is known.

"Questions are typically classified by the level of cognitive demand required to answer them." [33], Bloom's taxonomy [34] is the best-known system for categorizing the cognitive level of questions. In this activity, six levels of cognitive demand move from the lowest-order processes to the highest. Lower-order questions are asked to remember and comprehend material that was previously read or taught by the teacher. Higher-order questions are asked in order to use information previously read or taught to create or support an answer with logically reasoned evidence. Although they serve different purposes, both types of questions are useful for the teaching-learning process. A recent revision of Bloom's taxonomy done by Anderson Krathwohl in 2001 expresses the levels as verbs instead of nouns [35]. Table 2.1 presented the levels of revised Bloom Taxonomy in detail.

Table 2.1: Revised Bloom Taxonomy

High Order Thinking (HOT)	Creating	Compile information together in a different way by combining elements in a new pattern or proposing alternative solutions.
	Evaluating	Present and defend opinions by making judgments about information, validity of ideas, or quality of work based on a set of criteria.
	Analyzing	Examine and break information into parts by identifying motives or causes. Make inferences and find evidence to support generalizations.
	Applying	Solve problems to new situations by applying acquired knowledge, facts, techniques and rules in a different way.
Low Order Thinking (LOT)	Understanding	Demonstrate understanding of facts and ideas by organizing, comparing, translating, interpreting, giving descriptions, and stating main ideas.
	Remembering	Exhibit memory of previously learned material by recalling facts, terms, basic concepts, and answers.

Gallagher & Aschner [36] developed a different taxonomy which categorizes questions into four types:

- **Memory** questions focus on identifying, naming, defining, designating, and responding with yes or no. Keywords are who, what, where, when.
- **Convergent** questions focus on explaining, stating relationships, comparing, and contrasting. Keywords are why, how, in what way.
- **Divergent** questions focus on predicting, hypothesizing, inferring, and reconstructing. Keywords are imagine, suppose, predict, if...then..., how might, can you create, what are some possible consequences.
- **Evaluative** questions focus on valuing, defending, judging, and justifying choices. Keywords are defend, judge, justify, what do you think, what is your opinion.

The Graesser–Person taxonomy [27] classifies questions according to the nature of the information being sought in a good answer to the question. Table 2.2 lists and defines these categories. A scaling of depth outlined by the amount and complexity of content produced in a good answer to the question is used for the 16 question categories. In some of their analyses, simple shallow questions (categories 1–4), intermediate questions (5–8), and complex deep questions (9–16) were differentiated. The scale of depth is validated to the extent that it correlates significantly ($r = .60 \pm .05$) with both Mosenthal's scale of question depth [37] and the original Bloom taxonomy of cognitive difficulty [34].

Table 2.2: Question Taxonomy Proposed by Graesser and Person (1994)

Question category	**Generic question frames and examples**
1. Verification	Is X true or false? Did an event occur? Does a state exist?
2. Disjunctive	Is X, Y, or Z the case?
3. Concept completion	Who? What? When? Where?
4. Example	What is an example or instance of a category?
5. Feature specification	What qualitative properties does entity X have?
6. Quantification	What is the value of a quantitative variable? How much? How many?
7. Definition	What does X mean?
8. Comparison	How is X similar to Y? How is X different from Y?
9. Interpretation	What concept or claim can be inferred from a static or active pattern of data?
10. Causal antecedent	What state or event causally led to an event or state? Why did an event occur? Why does a state exist? How did an event occur? How did a state come to exist?
11. Causal consequence	What are the consequences of an event or state? What if X occurred? What if X did not occur?
12. Goal orientation	What are the motives or goals behind an agent's action? Why did an agent do some action?
13. Instrumental procedural	What plan or instrument allows an agent to accomplish a goal? How did an agent do some action?
14. Enablement	What object or resource allows an agent to accomplish a goal?
15. Expectation	Why did some expected event not occur? Why does some expected state not exist?
16. Judgmental	What value does the answerer place on an idea or advice? What do you think of X? How would you rate X?

In Mosenthal [37], a coding system was developed in order to estimate the level of questions' abstractness, which matches the depth of the question to some extent. The level of abstraction as suggested by Mosenthal [37] vary between most concrete, targeting explicit information, to intermediate levels which target information that could be explicit or not (such as Procedures, goals) and to abstract levels which target information meant to identify causes and effects, reasons and evidence. Similar to the taxonomy of the Graesser–Person scheme, the classification of Mosenthal focuses on the information desired from the answer. However, the "world knowledge" and "cognitive processes" are needed to generate answers to questions are not taken into consideration. Although both classification schemes combine possibly separable dimensions, both have a certain level of validity in estimating question quality for promoting learning.

In addition, Guthrie and Taboada [38] defined a questioning rubric for measuring student-generated questions quality for conceptual knowledge, which comprises four levels of question quality. These levels are:

- **Level 1:** Factual Information. In this level, questions are simple in form and request a simple answer, such as a single fact. Examples of this level of question are: How big are bats? Do sharks eat trash? How much do bears weigh?

- **Level 2:** Simple Description. In this level, questions request general information that denotes a link between concepts. The question can be simple, yet the answer may contain multiple facts and generalizations.
- **Level 3:** Complex Explanation. In this level, Questions request for an elaborated explanation about a specific aspect of concept with accompanying evidence.
- **Level 4:** Pattern of Relationships. In this level, questions display science knowledge coherently expressed to probe the interrelationship of concepts, they are a request for principled understanding with evidence for complex interactions among multiple concepts and possibly across concepts.

As explained in the previous section and this section, prior works have investigated the quality of the question. Researchers have supported students to generate questions by providing them with different types of scaffolding [39, 40, 41, 42], which are created based on the aforementioned criteria and/or the types of questions.

2.4 Function and Types of Scaffolding

Wood et al. [39] defined the term of "scaffolding" as an approach to help students reach pedagogical goals that might be difficult to achieve if they perform it alone. More particularly, scaffolding serves as a link to join the student's current abilities (actual development) and the goals they wish to achieve (potential development), which can later be discarded as the student becomes more competent in terms of skills.

In literature, diverse functions and types of scaffolding have been investigated from different viewpoints. Jackson et al. [40], suggested three scaffold designs, namely: supportive, reflective, and intrinsic scaffolding, each one performs different purposes in software systems:

- Supportive scaffolding: provides advice and support all along the process and ensures purposes such as guiding, coaching, and modeling. For example, including a built-in button such as "Show me an example" could help context-sensitive learners.
- Reflective scaffolding: uses forms and prompts to obtain statements from learners for planning, explaining, testing, and evaluating. This aims to help learners reflect on how they solve or conceptualize tasks.
- Intrinsic scaffolding: along with making a task less complex, this design uses visual support, such as maps and models, in helping learners reflect upon concepts.

Hannafin et al. [41] identified four kinds of scaffolding that can be used to foster students' learning in open-ended learning environments: conceptual, metacognitive, procedural, and strategic scaffolding:

- Conceptual scaffolding facilitates understanding the complex problem space currently under consideration or clarifying misconceptions through providing structural maps, content trees, or explicit hints to learners.

- Meta-cognitive scaffolding helps learners handle their individual thinking processes; learners are reminded to reflect upon their goals or propose self-regulatory strategies and related monitoring processes.
- Procedural scaffolding assists learners in using the available tools and resources, usually through the use of a 'balloon' or 'pop-up' help window.
- Strategic scaffolding offers alternative approaches or techniques for learning or suggestions for initial questions.

As for Ge and Land [42], they identified three kinds of scaffolding for question prompts: procedural, elaborative, and reflective:

- Procedural prompts guide learners to finish their specific tasks and facilitate them to spot and examine the important features of those tasks.
- Elaboration prompts support learners in articulating their thoughts, construct explanations, create justifications, and perform reasoning while using prompting questions such as: "Why is it important?" and "How does .... affect ...?"
- Reflective prompts facilitate learners' reflection and self-monitoring alongside the process by using embedded questions such as "What is our plan?" and "Have our goals changed?"

Based on the reviewed literature, we think that using procedural prompt techniques is the most suitable type of scaffolding for student-generated questioning since it guides students during the process of generating questions.

2.5 Techniques in Procedural Prompt

The procedural prompt is one of the types of scaffolding, where it is designed to help students' complete specific tasks such as writing [43] or problem-solving [44] by providing specific procedures or suggestions. Students can rely on these hints or suggestions until they are capable of performing it alone. A number of different procedural prompts were used in previous studies to help students learn how to generate questions.

According to Rosenshine [2], five techniques were commonly used in previous studies for the procedural prompt: (a) signal words. (b) generic question stems and generic questions. (c) the main idea of a passage. (d) question types. (e) story grammar categories. Each of these prompts was used in a way to help students to generate questions with or without the help of a teacher.

Signal words. A well-known and frequently used procedural prompt for helping students generate questions consists of first providing students with a list of signal words for starting questions, such as who, what, where, when, why, and how. Students are taught how to use these words as prompts for generating questions.

Generic question stems and generic questions. The second most frequently used procedural prompt was to provide students with generic questions or stems of generic questions. Students were given generic question stems in three studies by King [11, 29, 8], and specific generic questions in the study by Weiner [45]. Following are examples of the generic question stems used in the studies by King [11, 29, 8]: "How are ... and ... alike?", "What is the main idea of ... ?", "What are the strengths and weaknesses of... ?", "How does ... affect... ?", "How does ... tie in with what we have learned before?", "How is ... related to ... ?", "What is a new example of... ?", "What conclusions can you draw about... ?", "Why is it important that ... ?"

In another study, Weiner [45] used specific generic questions; the following questions were provided: (1) How does this passage or chapter relate to what I already know about the topic? (2) What is the main idea of this passage or chapter? (3) What are the five important ideas that the author develops that relate to the main idea? (4) How does the author put the ideas in order? (5) What are the key vocabulary words? Do I know what they all mean? (6) What special things does the passage make me think about? (p. 5)

Main idea. In the third type of procedural prompt, students were taught or told to identify the main idea of a paragraph and then use the main idea to prompt the development of questions. Using a booklet, Dreher and Gambrell [46] taught this procedure to high school students and sixth-grade students; the instruction included the following suggestions:

1. Identify the main idea of each paragraph.
2. Form questions which ask for new examples of the main idea.
3. If it is difficult to ask for a new instance, then write a question about a concept in the paragraph in a paraphrased form.

In other studies, similar procedures were taught orally to learning disabled and regular education junior high school students [47], to above-average junior high school students [48], and to fourth and sixth-grade students [49].

Question types. Based on Raphael and Pearson [50], in this procedural prompt, questions are divided into three types that can be differentiated according to the relationship between the question, its answer, and the amount of cognitive processes needed to proceed from question to answer. The three types of questions are (a) a question whose answer can be found in a single sentence. (b) a question that requires integrating two or more sentences of text, and (c) a question whose answer cannot be found in the text but rather requires that readers use their schema or background knowledge. In Raphael and Pearson's study, students were taught to identify the type of question they were being asked and to decide upon the appropriate cognitive process needed to answer the question. In their studies Dermody [51], Labercane Battle [52], and Smith [53], students were taught to use this classification to generate questions.

Story grammar categories. In two studies [54, 55], students were taught to use "story grammar" to help understand the narrative material they were reading. In the study [54], story grammar was used to help fourth and fifth-grade students in which consisting of four elements: (a) setting. (b) main character. (c) character's goal, and (d) obstacles. Students were taught to generate questions that focused on each element. For example, for the character element, they were taught that the set of possible questions included the following: "Who is the leading character?" "What action does the character initiate?" "What do you learn about the character from this action?"

2.6 Collaborative Question Generation and Peer Learning

Collaborative question generation boosts confidence and critical thinking abilities, it highlights various problem-solving solutions, supports imagination, encourages students, promotes rapid peer-assessment and enhances deep learning [56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72]. It also helps in the creation of question stems, the correction of wrong choices, justifying that each choice requires critical thinking and evaluating that result in deep learning [59, 73, 67]. Generating multiple-choice questions also involves metacognitive thinking to understand why other students may choose one of the right or wrong answers [73]. While developing a false choice, students notice errors and invent their own solutions to these errors, which contributes to their deeper learning [59].

PeerWise [74, 56] is a popular online system for collaborative multiple-choice question-generation. This system offers students a learning environment where they can generate questions and share them with their classmates for self-testing. Question-generation through PeerWise is found helpful in a large spectrum of disciplines [57] [59]. Replacing weekly course assignments with question generation also showed a significant increase in final exam grades [64]. Moreover, question generation is shown to improve not only exam performance but also students' engagement in more learning activities [62, 59].

Numerous studies have attempted to involve students in generating questions collaboratively through peer review during the generation process. Peer assessment has been shown to facilitate and enrich formative assessment. Piaget [75] acknowledges that opinions or arguments between individuals are the catalysts for meta-cognitive reasoning to draw convincing conclusions. Evaluating peers generated questions is helpful for their peers, where the peers are uncertain about the correctness of the questions generated by their peers and to participate in metacognitive thinking to clear up any uncertainties. PeerWise used peer assessment as a way to evaluate and examine the quality of the questions while providing constructive feedback on the generated questions.

Through the feedback given by their peers, students can refine their own questions, which may lead to better quality.

2.7 Question Refinement

The purpose of question refinement is to improve the quality of a question in a certain parameter, different researches using different approaches have been conducted in order to investigate the improvement of questions' quality [11, 15, 16, 17, 19, 18, 76].

King et al. [11] have applied question refinement through a continuous process of self-questioning and work in groups to ask the question and answer each other's questions, in hopes that students can improve their questioning and answering.

Moreover, studies like Denny et al. [15] and Yu [16] have also demonstrated the question refinement by asking students to evaluate multiple-choice questions created by others based on the quality of the question, distractors and the explanation why the answer is correct, and allow students to revise their questions. Yu [16] stated that "writing comments" was perceived as not only helping the question-authors to "refine their work," but also as a benefit to assessors because it induced them to "think critically" and to "learn to respond in a socially acceptable way for the purpose of communication.", and the "receiving comments" helped students "detect and correct incomplete ideas or misconceptions" and "improve one's questions". In addition, Yeckehzaare et al. [17] used this question refinement technique in a different approach in which the students can modify each other's questions and claim ownership of the modified question by adding justification to why the previous question needed to be refined.

Based on the reviewed literature, it appears that researchers in the education field have used the question refinement technique implicitly by compromising the assessment in learning with peer-assessment that asking students to evaluate other's questions and hoping students would use the feedback to improve their own questions.

Researchers in machine learning have also tried to improve the quality of a question in terms of clarity, spelling, and grammar. For instance, Elgohary et al. [18] constructed a dataset that contains standalone questions by improving the clarity of the questions by adding context to the questions such as "What happened in 1983?" to "What happened to Anna Vissi in 1983?". Furthermore, Liu et al. [19] proposed a unified model called Qrefine to refine ill-formed questions to well-formed questions.

While previously mentioned researches focused on improving questions' quality through using question refinement technique, a research done by Podgorny et al. [76] has used this technique to make the question simpler in order to make it easy to answer (less cognitive demand), such as refining "why" questions to "close-ended" questions, for example: "Why is my debit card being declined" to "My Debit card has been declined. Is there something I need to do to make it work?", which downgraded the question's quality.

However, to the best of the author's knowledge and the observation from the previous works, we conclude that there has not been any research focusing on the process of question refinement to have a better quality in terms of cognitive demand.

2.8 Activities in Creative Processes

According to Buxton [77] in Perteneder et al. [14], creative processes might be best to involves three separate activities, namely, gathering of inspiration, generation of ideas, and refinement of these ideas, as shown in the following Figure 2.1:

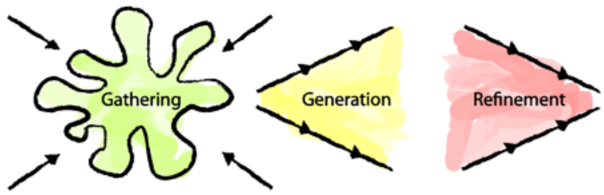


Figure 2.1: The three primary activities of creative processes

Gathering occurs everywhere in an amorphic, amebic way, including collecting artifacts, objects, as well as ideas. An example of gathering inspiration, designer studio walls are frequently decorated with these artifacts as methods for motivation to influence and stimulate the procedure for both individual and group-based. Nowadays, this activity of gathering inspiration can be done both digitally, such as browsing on the web and physically in the world we live in.

Perteneder et al. [14] described several ways to support gathering, hunting, and collecting the inspirational ideas: (1) Ubiquitous Ways of Input and (2) Synchronous and Asynchronous Operations. *Ubiquitous Ways of Input* is basically utilizing the wide variety of devices and media types as input. For example, cameras and smartphones can be used to take a snapshot to capture an idea that can then become content. *Synchronous and Asynchronous Operations* is basically allowing the input of ideas in different situations without limiting the type of content and providing asynchronous and synchronous activity.

As for ideas *generation*, brainstorming plays an essential role in this activity, where the goal is to be deliberately expansive. As the traditional definition of brainstorming [78], people are to be uncritical and deliberately unconventional. There are at least three major variations of brainstorming (verbal, nominal, and electronic), each with its own strengths and weaknesses.

In order to support various brainstorming techniques, Perteneder et al. [14] suggested providing a flexibility that supports more than one of these variations, allowing people to adopt the style that suits their needs. In addition, the environment should be given enough flexibility so that people can change the way they view the process. For example, a shared workspace could promote group well-being, which might allow people to overcome evaluation apprehension by anonymous contribution.

While ideas *refinement* implies working with the ideas generated by organizing, structuring, and selecting, an example of common practices is critique. Perteneder et al. [14] described various tools to support idea refinement in the whole process. As mentioned above, while idea gathering and generating is about expansive thinking. Refining ideas focus on manipulating ideas by reorganizing, structuring, altering, and or removing information. Refinement mechanisms include interaction techniques that can provide a way to rearrange content and visualize dependencies quickly. For example, clustering, grouping, consolidation, linking, coloring, and adjustment of size.

Chapter 3

Question Refinement Strategy

To answer our research question on how student's posed question can change or evolve, we proposed question refinement strategy, a learning strategy (Figure 3.1) adapted from the activities in creative processes (Figure 2.1) and the student-generated questioning. In the Gathering Stage, students are given a guidance on how to create questions and they have to watch a lecture video as the learning material. In the Generation Stage, students are asked to generates question related to the learning materials. In the Refinement Stage, students are asked to refine the previous created questions. A detailed explanation of these stages from our proposed learning strategy will be described in the following.

In order for students to generate better quality questions, our proposed strategy followed the suggestion from Chin [13], where students are asked to first generate their questions individually, and then redefine the questions in groups so that they would be more focused and investigable (doable, interesting, practical, and unknown).

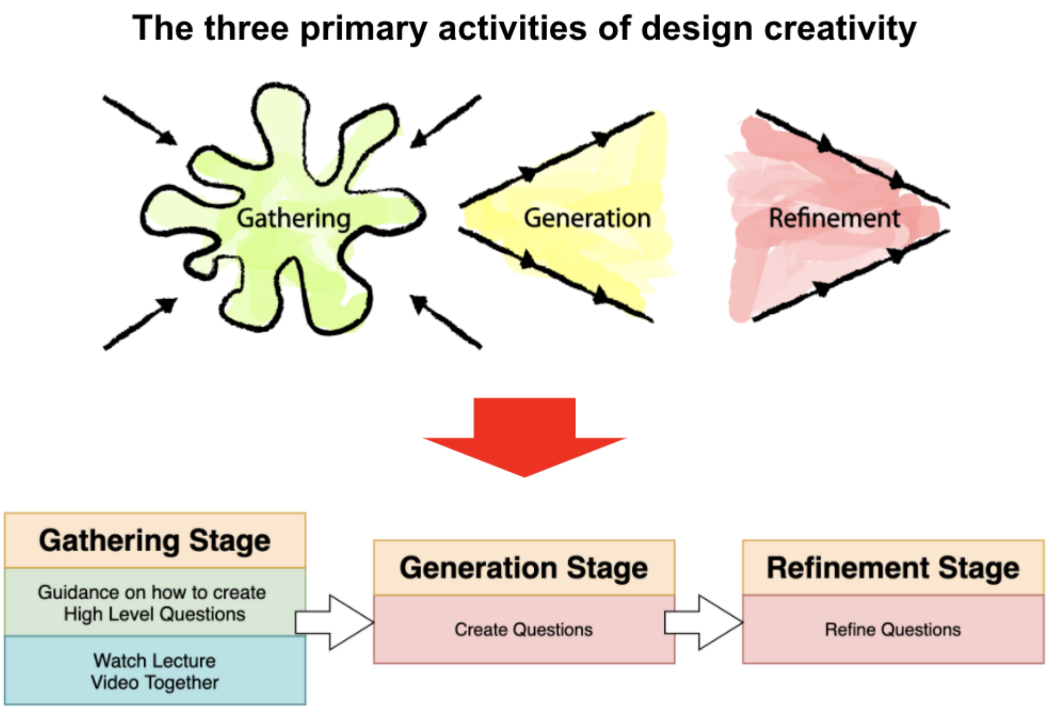


Figure 3.1: Our proposed learning strategy adapted from the three activities in creative processes

3.1 Gathering Stage

In the *gathering stage*, students are provided with guidance on how to create high-level questions as a scaffold in order for them to be able to generate and refine questions. Students are also provided a lecture video for the learning material, which will be used for the content of the questions.

Educators widely use Bloom taxonomy for measuring cognitive demand. However, Papinczak et al. [9] reported that medical students had difficulty creating questions using Bloom Taxonomy. Besides, based on our observation, the classification between each level in Bloom Taxonomy is hard to carry out. Moreover, Aflalo [79] also stated that although education college students were familiar with the Bloom Taxonomy, due to their little experience, students still had difficulty classifying the questions. Thus the guidance material is derived from the questioning rubric by Taboada and Guthrie [38] as in our observation it provides clearer distinction between quality levels compared to the other taxonomies. In addition, based on our pilot study the questioning rubric also considered easy to understand by the participants. This rubric has also been employed in Guthrie and his colleagues studies [80, 38, 81].

Furthermore, based on the questioning rubric's level, a procedural prompt is chosen to be the type of scaffolding of the guidance, as it may help the students to identify and analyze the important features of the levels in the questioning rubric and guides them to complete the tasks (generate and refine questions) [42]. Moreover, question type is the technique chosen as for the procedural prompt so the students are able to recognize the types/classifications of the question they generated according to the appropriate cognitive process which is needed to answer the question [53, 79, 50, 52, 51, 13].

The guidance was presented in two forms, presentation slides, and handout document. The presentation slides are used to introduce and familiarize the students with the concept of high-level questions in terms of question quality. King [8] proved that students who received guidance in generating questions tend to ask challenging questions which requires them to think critically about the material and generate good quality answers as they need to process the ideas more deeply and build broader cognitive networks that connect the new ideas and the previous knowledge they already have.

The handout document is an additional support for students to generate and refine questions, as a previous study [9] reported, students required continuous support to refine questions.

A lecture is given in a video format as it is proved that using videos can lead to better learning outcomes [82]. In addition, videos can help learners by visualizing how something works [83] and show information in detail, which are difficult to describe by text or static image [84, 85]. Furthermore, videos can also attract learner's attention, which motivates them and engages them to increase their collaboration [86].

3.2 Generation Stage

In the generation stage, students are asked to create questions as many as they can in limited time based on the video's learning content. According to Rosenshine [2], question generation is considered as a cognitive strategy and a metacognitive strategy as the process of asking questions enhances comprehension through a focus on main ideas (content) and also checks understanding to determine whether the content is learned.

Furthermore, while generating the questions, students are allowed to open any resources they are provided for creating the questions, including the lecture video and the handout guidance on how to create a high-level question that was given after the gathering stage. Making these materials available helps students to generate the questions since creating a question required to have some prior knowledge about the topic [87, 7, 68].

In addition, in this stage, students were given a time limit as a constraint, as the longer the time length, the less motivated students might get. According to Li [88], students' motivation to create questions significantly affects their learning [88], and this motivation typically declines over time [89, 90]. Several pilots have been conducted to decide the time length, and we found out that 10 minutes is adequate for most participants to create 4 to 6 questions. However, when students are done with generating questions, they can finish before the time limit ends and move on to the next stage.

Following the idea from Perteneder et al. [14] that, rather than restricting people to one particular style that suggests providing flexible facilities that support more than one of these variations, thereby allowing people to choose the style that suits their needs, a free format question is used in this stage of our study in which:

1. Question can be express not only in text format but also in other form such drawing and others, and
2. Question can be in both close and open-ended questions.

3.3 Refinement Stage

In the *refinement stage*, students are actively refining questions in 1 refinement activity for individuals or 3 refinement activities for groups. There will be three refinement activities in this stage and it is explained as follows: (1) Collaborative Questions Refinement, (2) Refining Own Original Questions, and (3) Collaborative Student Posed Refining Question.

Collaborative Questions Refinement - this activity is where the students refine others' previously created questions. For example, after student A generates a question in their own workspace, a copy of the workspace is shared with student B who is going to refine the question created by student A.

Refining Own Original Questions - this activity is where the students refine their own questions. For example, after student A generates a question in their own workspace (online board), student A is asked to refine their own question in their own original workspace.

Collaborative Student Posed Refining Question - this activity is where the students refine their own question after being improved by the others. For example, after student B refined student A's question in a copy of student A's workspace, the copy of student A's workspace is shared with student A in order to refine the question refined by student B.

Like in the generation stage, students are allowed to access the learning materials and the handout guidance on how to create a high-level question in order to refine the questions. The learning materials can be used for deepening students' knowledge about the topic. The handout can be used as a reference for improving the questions, for instance: an example of an improvement from descriptive questions such as "How do birds fly?" to complex questions "How the wings on birds can make them fly in the air?". A time limit is also set for this stage with a maximum duration of 5 minutes. The experiment trails also decided this time-length.

Chapter 4

Initial Study

To evaluate the feasibility of the proposed learning strategy, a small-scale initial study was conducted. This study has been approved by the University of Tsukuba Library Information Media Research Ethics Review Committee with the approval number 20-24. The study used a between-subject design with 3 conditions to investigate the process of collaborative (two students and three students) and individual activities on refining question to produce better quality in one particular video lecture:

1. **Individual Question Refinement** – Students will create and refine their own questions individually.
2. **Pair-student Question Refinement** - Two students will create questions, refine each other's questions, and refine their own questions after being refined by their partner.
3. **Triad-student Question Refinement** - Three students will create questions, refine their partner's questions one by one, and refine their own questions after being refined by their partners.

In this study, students watched a lecture video together in shared screen through Zoom meeting and continued with question posing activity using our question posing form (QPR form) on an online board. Question refinements are done by handing over the online board to the other students, where each student needs to refine the question posed or refined by the previous student.

Our main objective in this study is to investigate the improvement level of refining questions in term of quality between individual and collaborative work. Below, we will explain the detail on this initial study from the equipment and tools to the procedure of this study.

4.1 Experiment Setup

The experiment will be conducted in a full online setting without any face-to-face interaction to replicate the distance/remote learning environment, where the students are geographically separated from each other. Moreover, this also support with the current

COVID-19 pandemic condition that physical distancing is recommended. The experiment will be using a couple of off-the-shelf tools such as Zoom and AWW App. All the interactions between students and experimenters will be online to support the online setting. Thus, an online communication tools such as video conferencing and a shared working space, which replicates papers, will be needed to generate and refine questions in free format while the students are not in the same place.

One of the popular video conferencing applications is Zoom. Zoom is chosen because it is easy to use and provides decent video and audio quality. Furthermore, Zoom also provides a lot of collaboration tools such as screen-sharing, chat box, and etc.

A Web Whiteboard Application or AWW App is a well-known online browser-based board which provides tools to write text and draw object in a shared workspace. AWW APP is chosen in this study since it replicates paper. Moreover, the provided pen and the text tools are the primary tools for question generation and refinement activity where students create and refine questions.

In this study, Zoom was utilized as the communication tools between experimenters and students. While, AWW App was utilized for the question generation and refinement activity, where every student will be given one AWW App board as their original working space to pose questions and will be given an access to other students original or copy board according to the experiment conditions.

4.2 Equipment Tools

Here is the list of equipment and tools that we used in our initial study:

Figure 4.1: Zoom application

1. *Zoom application* (Figure 4.1) was used as a communication tools between experimenters and participants such as giving explanation and instruction, sharing links and documents, and for communication between participants such as discussion. In this study, the screen sharing feature in Zoom was utilized for playing an explanation video about the purpose and the procedure of this experiment, a demonstration video on how to fill the consent form, a guidance video on how to create high quality question, a training video on how to use the other tool (AWW App board),

and the video lecture to the participants. Moreover, the chat feature in Zoom was also utilized to share the links and documents to the participants. Furthermore, using the recording feature, we recorded all of the communication in the experiment session for further analysis. In this study, the premium version of Zoom was used since the estimated time for one experiment session is more than time limit of the free version.

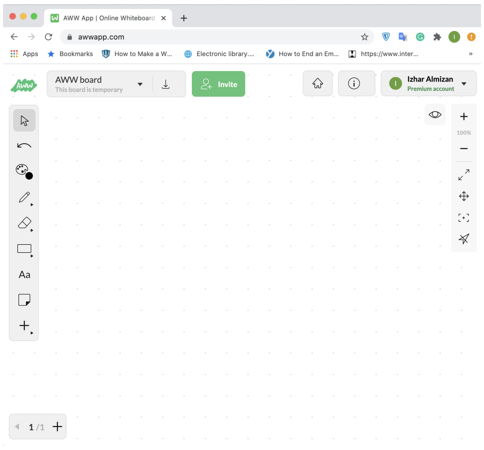


Figure 4.2: AWW application accessed through a web

2. *A Web Whiteboard Application or AWW App* accessed through a web (Figure 4.2) was used for the online collaborative question generation and refinement activity. This online board (workspace) was utilized as a replacement of an actual paper. Each participant will have one board as their original board (workspace).

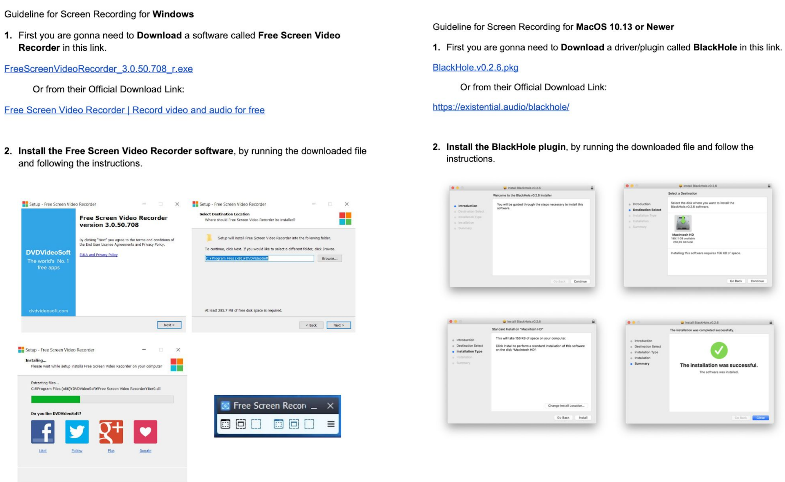


Figure 4.3: Guidelines for screen recording.

3. *Guidelines for screen recording* (Figure 4.3). We provide a step-to-step guidelines on how to use screen recording in both MacOS and Windows. In the guidelines we ask the participants to install an additional plugin (BlackHole for MacOS) or software (Free Screen Video Recorder for Windows) to be able to record their screen entirely and also to record the audio coming from the computer not only the audio coming from the microphone. Both plugin and software that we recommended to install are coming from developers that are trusted and free from spyware or other adware. We distribute the guidelines using Google Drive Link, so all the participant can access it easily.

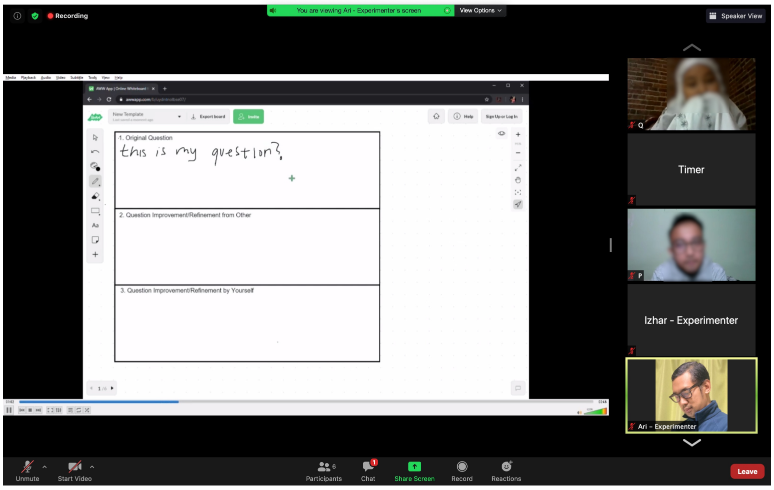


Figure 4.4: Guidelines for using the AWW App through Zoom

4. *Guidelines for using the AWW App* (Figure 4.4). We provide a video guidance or video training about how to use the AWW App through screen sharing from Zoom. In the training we teach the participants to be able to use the primary tools needed for this study and how they can use it to complete the given task, so later on they do not have any trouble when performing the task. For Example, in the left menu or toolbars, there are 9 tools (select, undo, color settings, pen, eraser, simple object such as rectangle and ellipse, text, post it, and upload) that participants use to edit and manipulate the AWW board, however we only focus on two tools. The first tool is the pen tool which the participants can use to write or draw in the board freely, and the second tool is the text tool, which the participant can use to write something in the board, we recommend the participants to use this text feature to write if they are having difficulty to write using the pen tool.

5. *Online consent form* (Figure 4.5) is used in order to provide evidence that the participants gave their consent that he or she agrees to the procedure and is aware of any risks that might be involved in the experiment. We provide the consent form through AWW App, so the participants can easily access it and download the written form as a PDF file for their own copy.

6. *Guidelines on how to create high-quality question* (see Section 4.2.1) to support the participants in question refinement activity.

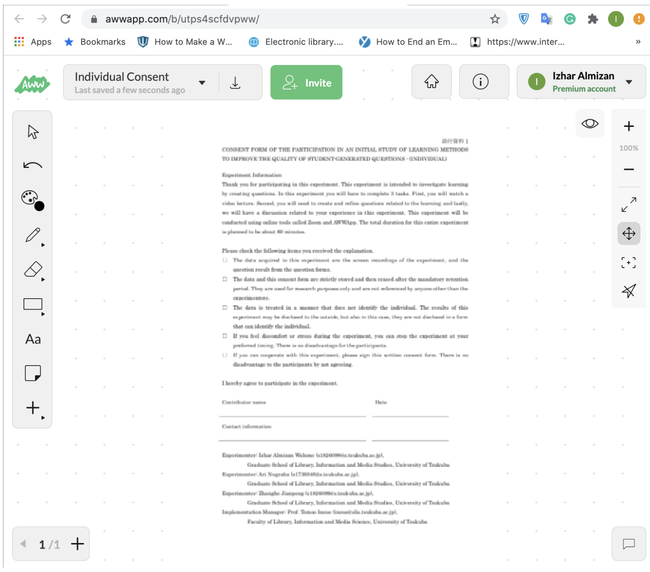


Figure 4.5: Online consent form through AWW App

7. *One lecture video* (see Section 4.2.2), for the learning material.
8. *Question Posing and Refinement (QPR) form* (see Section 4.2.3), for student to generate the question and a way to collect the generated questions.
9. *Online Questionnaire* (see Section 4.2.4) for measuring the learning experience from each participant.

4.2.1 Guidelines on How to Create High-Quality Questions

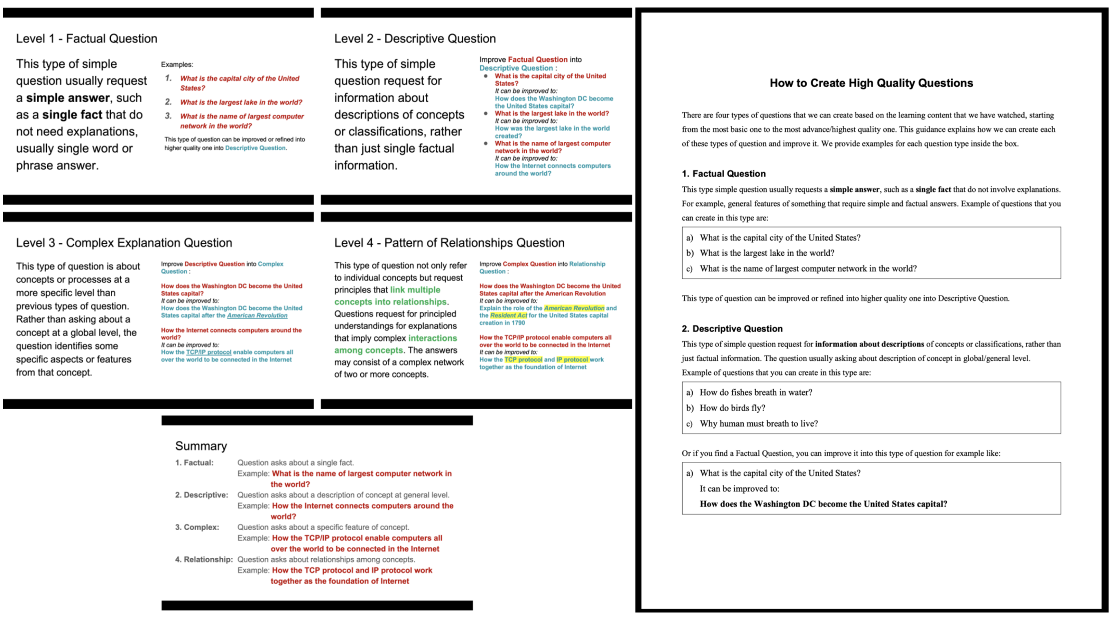


Figure 4.6: Snapshots of the guidance video on how to create high-level questions. Left: Video presentation slide, Right: Handout document

Guidelines or guidance on how to create high-quality questions were provided to every participant in this initial study. The guidance material was based on the questioning rubric by Taboada and Guthrie [38]. This guidance will be provided to the participants in two forms (see Figure 4.6), video presentation slide and handout documents. The video presentation slide will be given through a shared screen in Zoom and the handout document will be provided through a link to the document in Google Drive. These presentation slides and handout document will be shared in the training phase before the participants start performing their tasks.

4.2.2 Learning Material

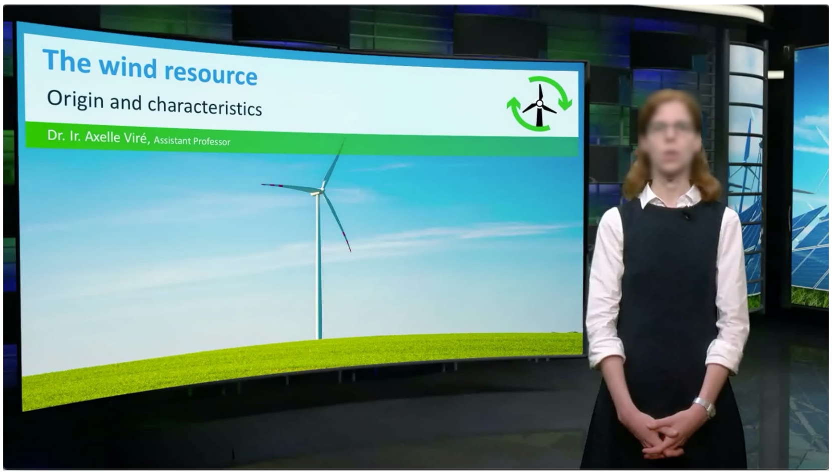


Figure 4.7: Snapshot of the lecture video (URL https://youtu.be/5peZJxzbMwE)

In this initial study, we defined 5 criteria to choose the lecture video as the learning material: (1) The format of video should be a talking-head style, where the presenter and the presentation slides appear side by side in the video as it reported to have better engagement [91], (2) The video duration should be around 0-9 minutes [91], (3) The subject/topic should not be popular in video sharing platform and it should be an introductory video, (4) The lecture should be given in English and the presenter articulates clearly, and lastly,(5) The video should be available for free usage.

We found an online lecture video that is matched with the above criteria. The video was retrieved from the "EnergyX: Sustainable Energy: Design A Renewable Future" YouTube Channel with a title of "The wind resource - Sustainable Energy - TU Delft", (URL https://youtu.be/5peZJxzbMwE). It was presented by Dr. Ir. Axelle Viré, an assistant professor in Delft University of Technology. Our other considerations from choosing the video is because the topic is relatively easy to understand for the general public since it gives an introduction to the topic explained.

Moreover, the video was edited to have a video subtitle, where it can help the participants easily understand what the lecturer is saying entirely. In order to distribute easily to each participant, the video was uploaded to a shared folder in Google Drive and later on the link to that folder was distributed through the Zoom Chat box. In addition, if the students are having difficulty with accessing the video through Google Drive, then we will also share the original YouTube link to the participants.

4.2.3 Question Posing and Refinement Form

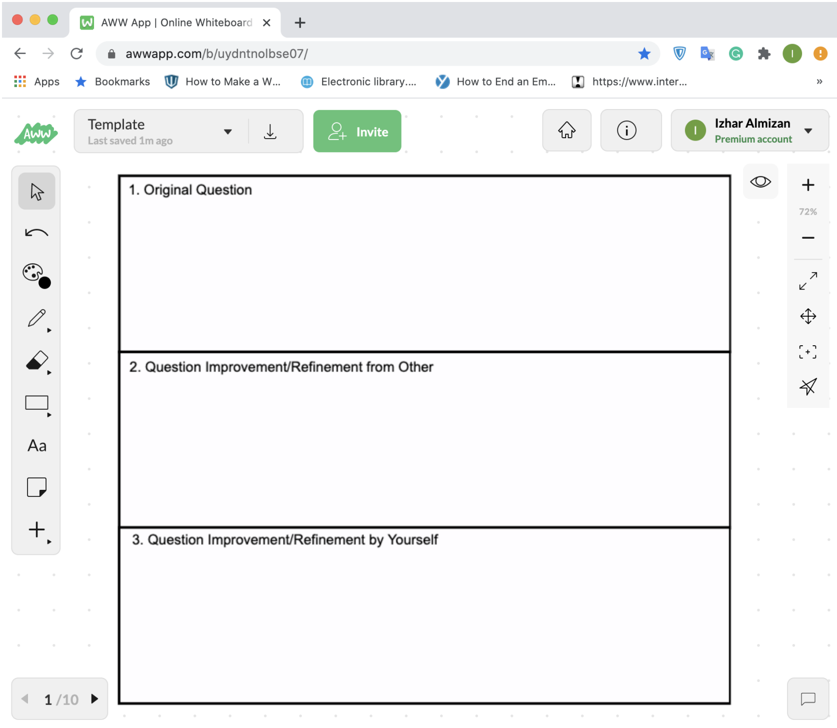


Figure 4.8: Question Posing and Refinement (QPR) form provided through AWW App.

The Question Posing and Refinement (QPR) form is designed with 3 boxes as illustrated in Figure 4.8, where each box have different purposes. The first box is used for the students to create their original questions and if they know the answer to the question, they could also write it in the same box. The second box is used for other students to create a refined question according to the original question in the first box. And lastly, the third box is used for students to generate a refined question based on the question that have been refined in the second box or the original question in the first box. One form is only for 1 original question, thus if the students want to create 2 questions, then they will need 2 different QPR forms.

Moreover, each box in the form was designed to have a large space for students to pose a question and not restricting the format of the questions that the students can create. Students can create questions using text as well as using drawings or pictures. All the conditions in the experiment will be using this QPR form as a working space for creating and refining questions through AWW App.

4.2.4 Online Questionnaire

The questionnaire was designed to measure the participants subjective rating on their learning experience based on learning experience questionnaire from [92, 93]. The items for questionnaire are shown in Table 4.1. The items will be rated using 7 points Likert scale.

Table 4.1: Questionnaire items

No.	Items
	Emotional Aspect
1	I felt engaged in this question creation and improvement activity
2	This activity was meaningful
3	I felt this activity was challenging in a stimulating way
4	This activity was fun learning experiences
5	This activity was too difficult
	Cognitive Aspect
6	I was able to learn many new things while doing questions improvement
7	I was able to get many new ideas and inspirations while doing questions improvement
8	This activity helped me to understand the key concepts in the video
9	This activity helped me to have deeper understanding to the concepts introduced in the video
10	This activity helped me to reflect my understanding to the learning content in the video
11	This activity helped me to memorize the learning content in the video
12	This activity inspired my interests towards the learning content in the video
	Collaborative Aspect
13	I prefer doing questions improvement collaboratively with friends
14	I prefer doing questions improvement individually
15	Having feedbacks is very useful to improve my own questions
	Teaching Material Aspect
16	The length/duration of the lecture video was good for learning
17	The length/duration of the lecture video was enough for me making questions
18	The learning content in the video was interesting
19	The learning content in the video increased my curiosity to the topic
20	The learning content in the video was too difficult for me
	Workspace Aspect
21	The workspace (online board) was easy to use for making and refining questions

4.3 Participants

In this initial study, 18 undergraduate universities students were recruited. Students were recruited through social network services such as Line, and WhatsApp. All the recruited students were unfamiliar with the topic in the lecture video. In addition, all the students are non-native English speakers. The students are divided into 11 groups (6 individual, 3 pair, and 2 triad) to balance the number of students in every condition.

4.4 Procedure

As it was mentioned above, this initial study was conducted in an online setting, where all the communication and work carried through Zoom and AWW App. The online meeting was scheduled using Zoom and the guidelines for screen recording were shared to the participants at least 1 day before the actual experiment.

The experiment was conducted in three stages for each experiment session as can be seen on Figure 4.9. The detailed procedure is described as follows.

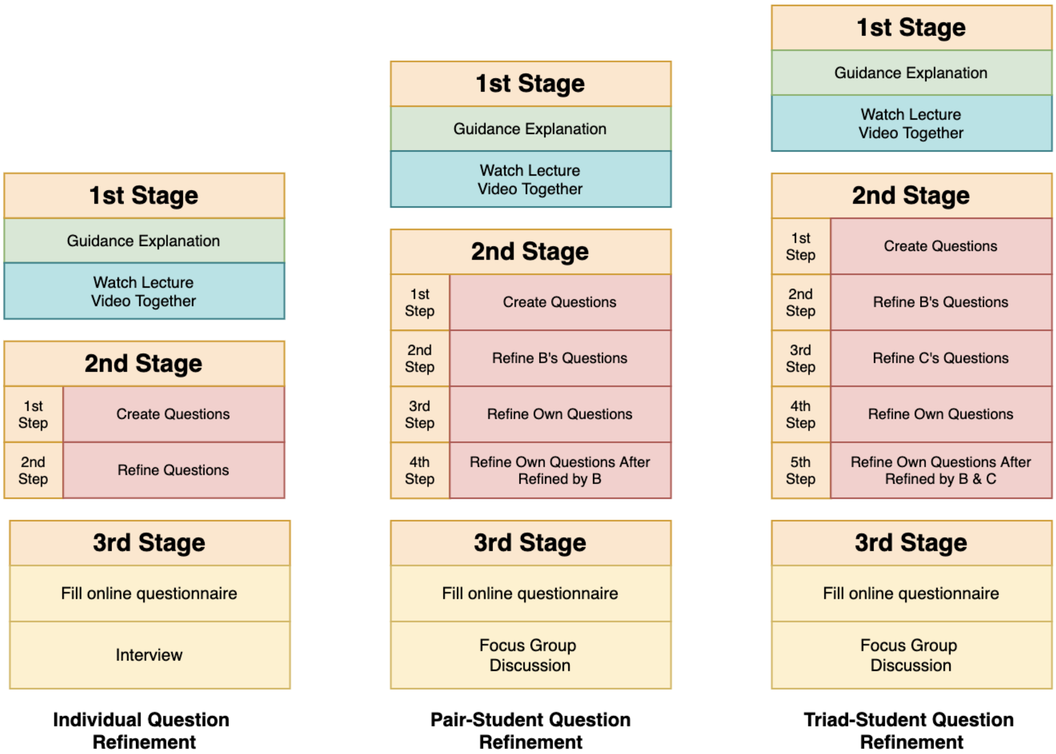


Figure 4.9: Experiment tasks stages from participant A's perspective

1. Ensure that the systems and tools being evaluated are functioning properly, the related instruments are ready for the experiment. Especially for the screen recording from participant's side, we will provide a guideline to the participants on how to record their PC screen prior to the experiment session. Written instruction also will be given related to the installation of additional screen recording software or plugin.

2. Gather participants for online meetings at the scheduled time.

3. Assign the participants to a specific experimental condition according to the predefined participants order.

4. Let the participants join the Zoom meeting and greet the participants.

   (a) In order to provide a safe identity for the participants, an alias names (P, Q, or R) were given to each of them, to be used for whole experiment.

5. Introduce the purpose of the study and the procedures to the participants.

   (a) An explanation video is given to the participants that explained about the purpose of the experiment and the tasks that they are needed to do in the experiment.

   (b) Participants are also notified that they are not allowed to do other activities during experiment which may disturb the experiment and have to record their own PC screen.

6. Get the consent of the participants.

(a) The consent form is provided using the AWW App, thus the participants can fill and sign the form easily.

(b) In order to guide the participants to fill the form, a demonstration video on how to fill the consent form in the AWW App is presented to the participants before they are asked to fill the form.

7. Participants complete the first stage in Figure 4.9.

    (a) Participants complete training phase.

        i. A video presentation slide was presented to the participants on how to create high-quality questions based on the questioning rubric by Taboada and Guthrie [38].

        ii. After the video is finished, the guidance handout document is shared to every participant through Zoom chat box so they could read it again if needed.

        iii. Participants are given another video that explained on how to use the AWW Application (online board) as their working space. Particularly the video explains on how to use the primary tools that the participants need to complete tasks. First the pen tool to write or draw something on the board. Second the text tool to write questions. And lastly, the page indicator to move around board's pages.

        iv. In order to train the students on how to improve a question and to make them familiarized with the online board, a test board with two example questions is provided to the students, so they can test using the AWW App and try to improve the provided questions to a better quality level based on the questioning rubric by Taboada and Guthrie [38].

    (b) Participants are exposed to watch a lecture video and allowed to take notes during watching the lecture video.

        - After the video is finished, the lecture video file is shared to every participant through the Zoom Chat box, in such manner, the participants can watch it again if needed.

8. Participants complete the second stage or Question Posing and Refinement Stage in Figure 4.9.

    (a) Participants are asked to start their screen video recording.

    (b) In this stage there are 4 activities:

        - **Generating Questions:** Participants are individually asked to create as many questions as possible and related to the lecture video content using his/her online board (AWW App) within a time limit of 10 minutes (based on the previous experiment trials, 10 minutes is adequate for most participants to create 4 to 6 questions).
        - **Collaborative Questions Refinement:** Participants will be asked to refine other's questions using a copy of other's online board (AWW App) within a time limit of 5 minutes per questions.

- **Refining Own Original Questions:** Participants are asked to refine his/her own questions using his/her original online board (AWW App) within a time limit of 5 minutes per questions.
- **Collaborative Student Posed Refining Question:** Participants will be asked to refine his own questions after being refined by others using the copy of his online board (AWW App) within a time limit of 5 minutes per questions.

(c) For participants in the individual condition, there are only 2 steps in this stage as presented in Figure 4.9. An example of the process from participant A's perspective is illustrated in Figure 4.10:

(1) **Generating Question** activity, where participants will create questions.

(2) **Refining Own Original Questions** activity, where participants will refine his/her own questions.

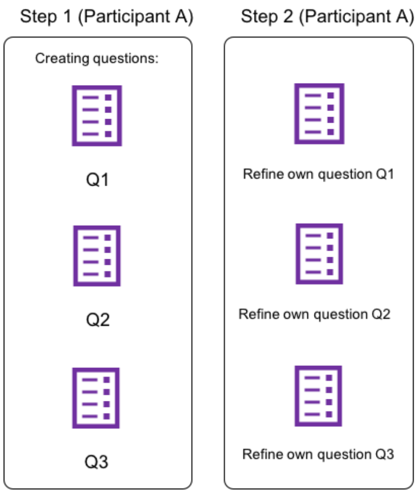


Figure 4.10: An example of stage 2 in the individual condition

(d) For participants (A & B) in the pair condition, there are 4 steps in this stage as presented in Figure 4.9. An example of the process from participant A's perspective is illustrated in Figure 4.11:

(1) **Generating Question** activity, where both participants A & B will create questions individually.

(2) **Collaborative Questions Refinement** activity, where:
- Participant A will refine B's questions, and
- Participant B will refine A's questions.

(3) **Refining Own Original Question** activity, where both participants A & B will refine their own questions.

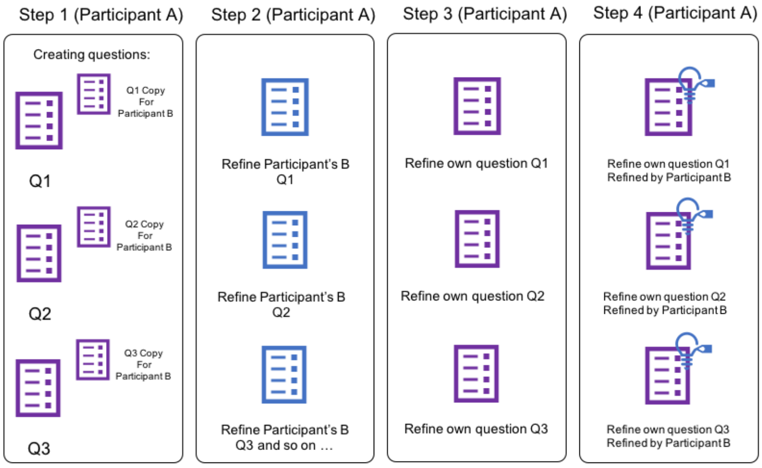


Figure 4.11: An example of stage 2 in the pair condition

(4) **Collaborative Student Posed Refining Question** activity, where:

- Participant A will refine his/her own questions after being refined by Participant B, and
- Participant B will refine his/her own questions after being refined by participant A.

(e) For participants (A, B, & C) in the triad condition, there are 5 steps in this stage as presented in Figure 4.9. An example of the process from participant A's perspective is illustrated in Figure 4.12:

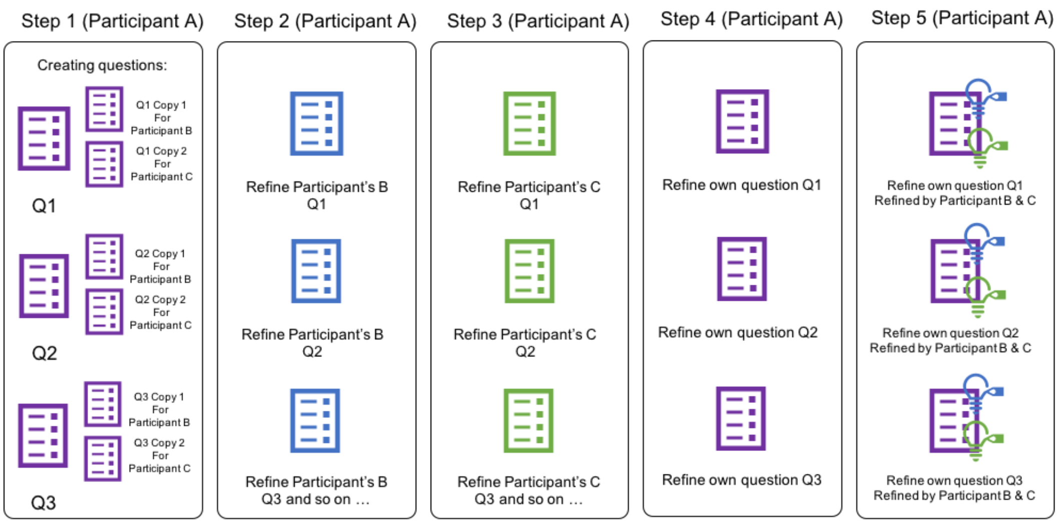


Figure 4.12: An example of stage 2 in the triad condition

(1) **Generating Question** activity, where all participants A, B & C will create questions individually.

(2) **Collaborative Questions Refinement** activity, where:

- Participant A will refine B's questions,
- Participant B will refine C's questions, and
- Participant C will refine A's questions.

(3) **Collaborative Questions Refinement** activity, where:

- Participant A will refine C's questions,
- Participant B will refine A's questions, and
- Participant C will refine B's questions.

(4) **Refining Own Original Question** activity, where all participants A, B & C will refine their own questions.

(5) **Collaborative Student Posed Refining Question** activity, where:

- Participant A will refine his/her own questions after being refined by participants B & C,
- Participant B will refine his/her own questions after being refined by participant A & C, and
- Participant C will refine his/her own questions after being refined by participant A & B.

(f) While creating or refining the questions, the participants are allowed to watch the video lecture again or read the guidance again with the provided links.

(g) The maximum duration for the 1st step (question generation) is 10 minutes and for the refinement steps is 5-minutes for each question generated. The refinement will be conducted one by one starting from the first question. For participants who do not have any more question to be refined he/she will be idle and must wait until other participants finish their refinement. Participants are allowed to skip refining the question by giving some mark such as - or ✓and write his/her reason to skip the question refinement.

(h) Since the experiment will be conducted online and in order to avoid confusion among participants in passing questions to the others, the experimenter will be passing the copy of the generated questions (Individual AWW App board) to the next person in predefined order for the refinement steps. Furthermore, this has proven to help participants stay focused on refining tasks during the trial experiment.

9. Participants are given an online questionnaire to be filled.

10. Participants complete the last or third stage in 4.9 after being interviewed (for individual condition) or conducted a focus group discussion (for collaborative conditions) with the following questions:

    (a) Did you get any benefit from engaging in this question generation and refinement activities? If yes, in what way? If no, then why not? Please explain and elaborate as much as possible [38].

(b) What do you think about the learning content or the lecture video itself? Did you think it was good and easy to understand? Please explain and elaborate as much as possible?

(c) What kind of difficulty or challenge that you found during this activity? Please explain and elaborate as much as possible.

(d) What do you think about the workspace or the online board that we use for this activity? Was it easy to use? Did you found any difficulty when you use it?

Chapter 5

Analysis & Results

To understand the effect of collaborative and individual activities in question refinement, we measured 2 variables in this initial study:

- Question quality level for every refined question, and
- Student's Learning Experience

In this study, we did not measure the comprehension level of the students since the focus of our study is on the effect of refinement activity to the previously generated questions.

5.1 Data Analysis Procedure

5.1.1 Questions Quality Level Analysis

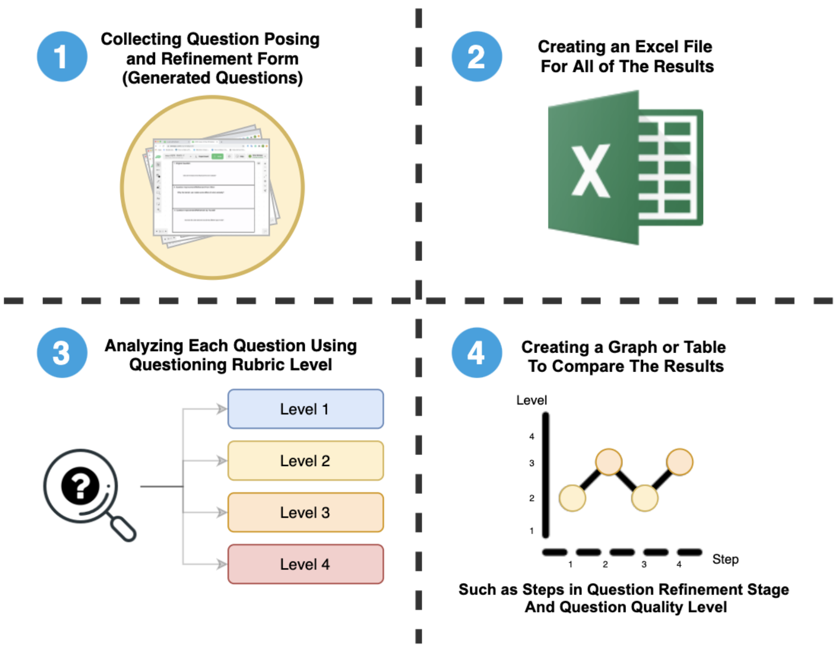
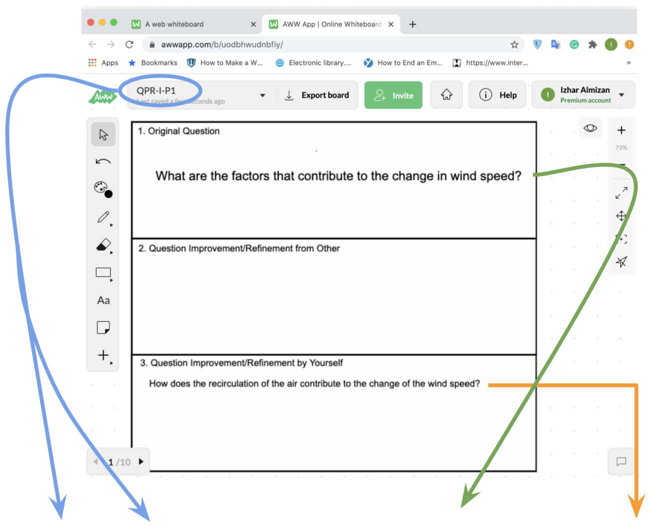


Figure 5.1: Question quality level analysis procedure

The collection of QPR forms data from the initial study were collected and analyzed in order to see whether different group treatment in term of the number of participants has significant differences on the question refinement quality, the increment in question quality level in each refinement steps will be observed and analyzed. One QPR form consists of 1 original question and 1 or 2 refined questions. The refined questions may also have a - or ✓symbol where the participant skipped not to refine the question.

A procedure comprising of 4 steps (see Figure 5.1) was conducted to analyze the quality level of the generated questions. First, all of the raw data (QPR forms) from the initial study were collected and organized according to the experimental conditions. We have gathered 168 QPR forms or 336 questions in total.



Group ID	Alias	Original Question & Answer	1st Refinement Step (by ownself)
P1	P	What are the factors that contribute to the change in wind speed?	How does the recirculation of the air contribute to the change of the wind speed?

Figure 5.2: Representation of step 2 in the QPR form data analysis procedure for individual condition

Second, the content of the QPR forms were transferred into an excel file (table form) with 3 sheets (Individual, Pair, and Triad). The 3 following columns are the same in every sheet:

- Group ID,
- Participant's Alias Name, and
- The Original Question.

In the "Individual" sheet, we have 3 more columns, which are the following: 1st Refinement Step (by oneself), Question Level of the Original Question, and Question Level of the 1st Refinement Question. Figure 5.2 represents how the columns filled from the QPR form from the individual condition.

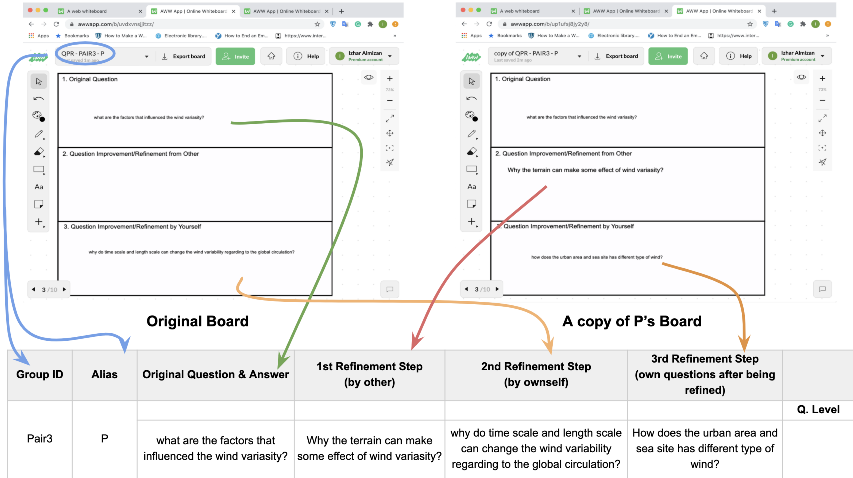


Figure 5.3: Representation of step 2 in the QPR form data analysis procedure for pair condition

In the "Pair" sheet, there will be 7 more columns as follows: 1st Refinement Step (by other), 2nd Refinement Step (by oneself), 3rd Refinement Step (for own questions after being refined), Question Level of the Original Question, Question Level of the 1st Refinement Question, Question Level of the 2nd Refinement Question, and Question Level of the 3rd Refinement Question. Figure 5.3 represents how the columns filled from the QPR form for the pair condition.

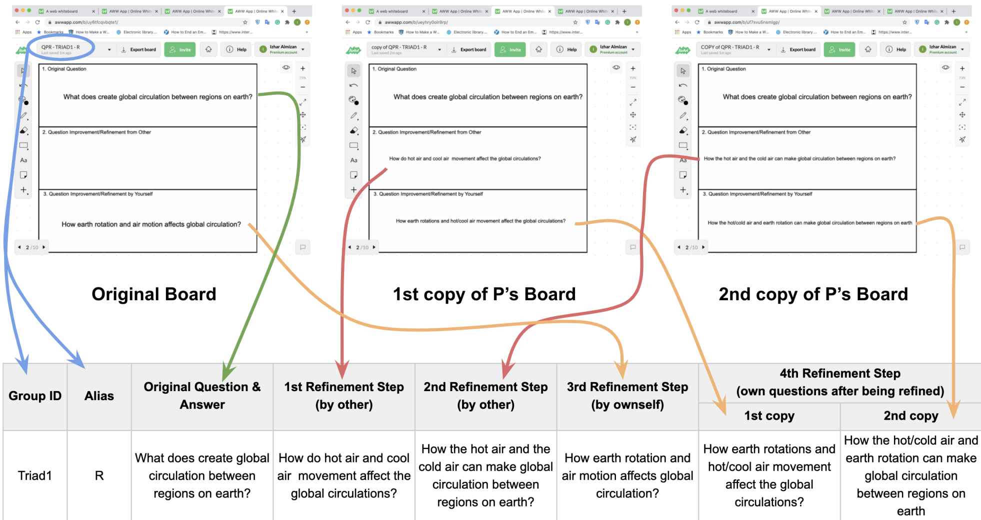


Figure 5.4: Representation of step 2 in the QPR form data analysis procedure for triad condition

In the "Triad" sheet, there will be 7 more columns: 1st Refinement Step (by other), 2nd Refinement Step (by other), 3rd Refinement Step (by oneself), 4th Refinement Step (for own questions after being refined), Question Level of the Original Question, Question Level of the 1st Refinement Question, Question Level of the 2nd Refinement Question, Question Level of the 3rd Refinement Question, and Question Level of the 4th Refinement Question. Figure 5.4 represents how the columns filled from the QPR form for the triad condition.

From the content of the QPR forms, half the columns in the excel file excluding the columns for question level could be filled. We repeated this step until all the QPR forms were transferred into the excel file.

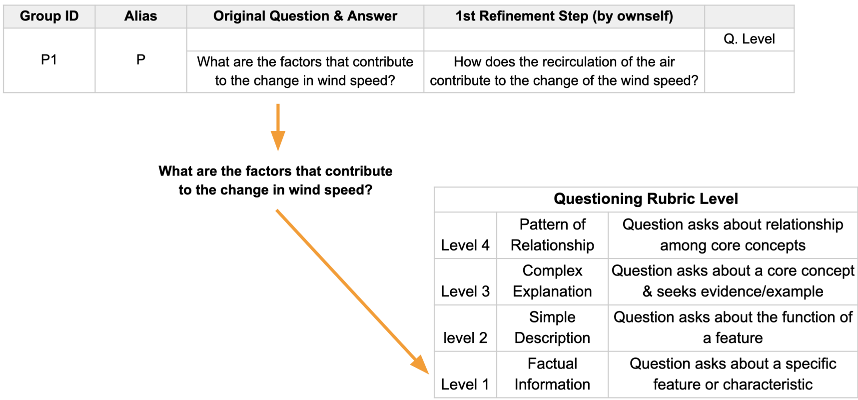


Figure 5.5: Representation of step 3 for normal questions in the QPR form data analysis procedure

Third, all questions (the original and the refinements) were analyzed by classifying the questions one by one to a specific level on the questioning rubric and filled in the question level columns. Although in Idea Generation field at least two experts are required in order to classify the quality of an idea. In this study, despite the fact that we are not an expert in the topic which is used in the learning material, we were able to determine the level of question quality by following the criteria in the rubric since the rubric already provides a clear distinction between each level. Moreover, when classifying the questions, the answer of the question was discarded since only the question was analyzed. In total we discarded 0 answers since no questions had answers. An example on classifying the quality of a question with the criteria of the four levels in the questioning rubric is illustrated in Figure 5.5. In this case, the original question *"What are the factors that contribute to the change in wind speed?"* matches with the criteria of level 2 where it is asking for simple description.

However, if the first refinement question in the third box of QPR form in Figure 5.2, the question will be *"How does the re-circulation of the air contribute to the change of the wind speed?"* and the quality level of the question does not match with the criteria of level 2 anymore since it is not a question asking about a simple description. Therefore, the question matches with the criteria of level 3 where the question is asking for complex explanation.

From the excel file, 32 refined questions have - or ✓symbol as its content which means the participants skipped to refine the question. In this case, these refined questions were classified as having the same quality level as the previous step's question. An example was shown in Figure 5.6, where in the third box of QPR form the refined question have the ✓symbol. In this case, the refined question was classified with having the same level as the question in the second box (level 4).

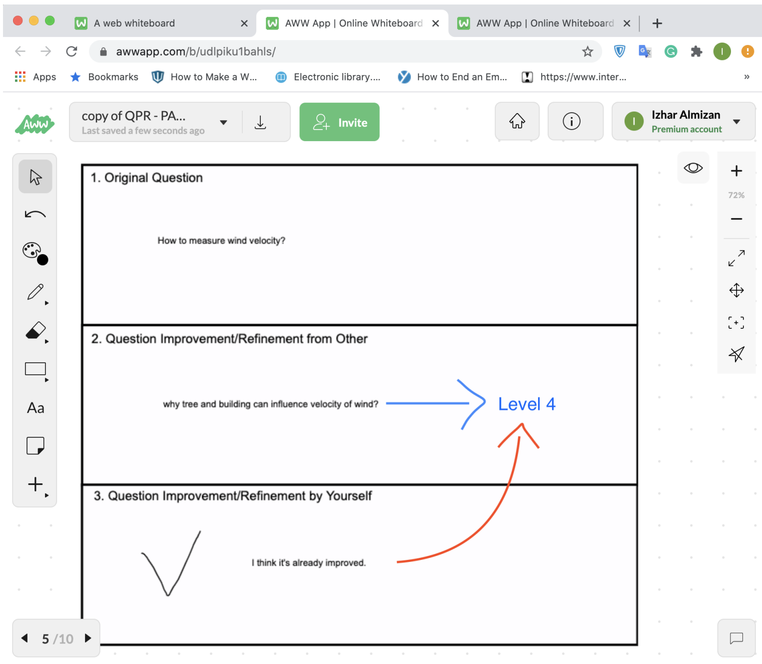


Figure 5.6: Representation of step 3 for skipped questions) in QPR form data analysis procedure

Lastly, in order to compare between experimental conditions (Individual, Pair & Triad), an analysis of variance (ANOVA) was conducted after all the 336 questions (including the skipped questions) were classified. The reason we used ANOVA for our analysis is because we are trying to compare the questions quality level generated in each experimental condition, including the refinement quality.

5.1.2 Student's Learning Experiences Analysis

In addition, an analysis on investigating the student's learning experiences and their impression to the lecture and online workspace was carried out. Questionnaire form with 21 items rated with 7 Likert scale were used to measure participants' learning experience. A two steps procedure were conducted to analyze the questionnaire results: collect and analyze. These 2 steps will be described as follows:

The first step in analyzing the student's experiences is by collecting the questionnaire results. Since Google Form was used as the platform for the online questionnaire, the retrieved results were already organized in an excel format, which helps analyzing it for the next step.

In the second step, the results were analyzed by using Statistical Package for the Social Sciences (SPSS) with a non-parametric test to investigate any significant differences between variables in the question items.

5.2 Result

5.2.1 Questions Quality Level

From the initial study, 336 questions which consist of 88 original questions and 248 refined questions were gathered. The results from each condition for each refinement steps will be presented within 3 sections, overall result, question refinement after making refinement, and question refinement after refinement feedback.

Overall Result

Table 5.1: Overall Question Quality Level

Experiment steps	**Individual**	**Pair**	**Triad**	**Total**
Generation Stage				
Level 1	6	10	6	22
Level 2	18	9	6	33
Level 3	1	9	4	14
Level 4	5	8	6	19
Total	30	36	22	88
Question Refinement to Other's Questions				
Level 1	-	1	1	2
Level 2	-	9	2	11
Level 3	-	14	20	34
Level 4	-	12	21	33
Total	-	36	44	80
Question Refinement after Making Refinement				
Level 1	2	0	0	2
Level 2	4	1	0	5
Level 3	14	23	9	46
Level 4	10	12	13	35
Total	30	36	22	88
Question Refinement after Seeing Refinement Feedback				
Level 1	-	1	0	1
Level 2	-	7	2	9
Level 3	-	15	13	28
Level 4	-	13	29	42
Total	-	36	44	80
Total from all of the steps	60	144	132	336

From the overall result, we observed how the evolution of the question quality from step to step, from the generation steps to the last step of the refinement stages. Table 5.1 present the results for the quality of all the generated question in this study.

Moreover, Figure 5.7 shown the average quality level generated per conditions in step by step from the whole experiment procedure. As we can see from the graph, the quality level in triad condition is steadily increasing, and the quality level in individual condition is significantly increasing, however it is still not as good as the pair and triad condition.

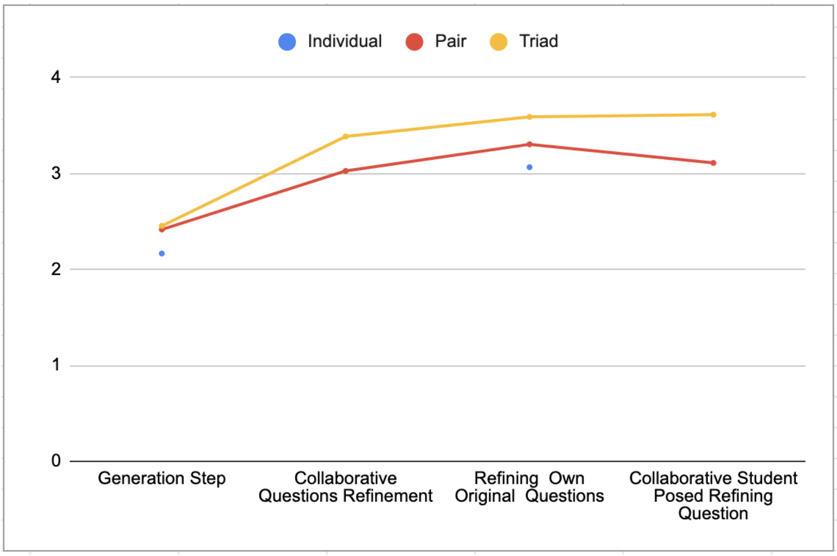


Figure 5.7: Question quality level changes from step to step

Question Refinement after Making Refinement

We observed how making refinement to the other participants' questions (step 3 in Pair condition and step 4 in Triad condition) could affected students in refining their own original questions. Our ANOVA result reported that there was a significant difference between Triad and Individual condition (p = .040). This result is represented in Figure 5.8. In addition, no significant difference found between the Pair and the Individual condition.

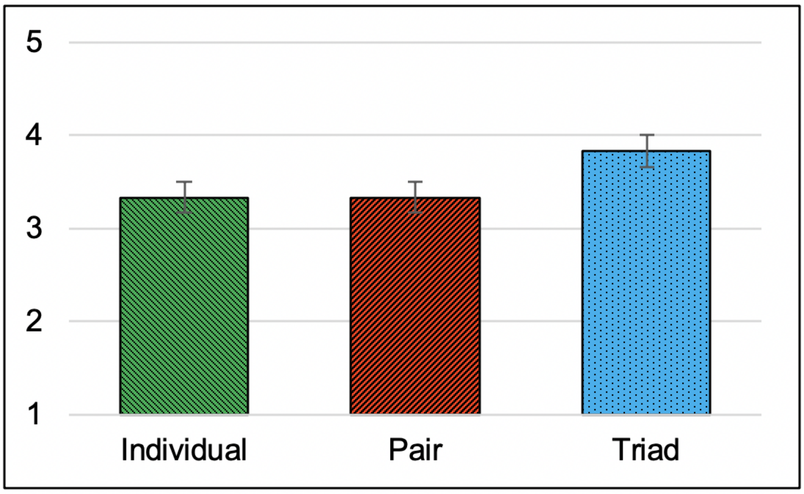


Figure 5.8: Question quality level after making refinement to others' questions

Question Quality after Seeing Refinement Feedback

To see how questions refinement from other participants can have an effect on question quality refinement to the original question poser (step 4 in Pair condition and step 5 in Triad condition), our ANOVA reported significant effect of condition to the question quality level after participants got question refinement from other participants, $F(2,15) = 4.113$, $p = .038$. This result is represented in Figure 5.9. In addition, Post-hoc LSD test revealed that the Triad condition was superior compared to the Pair condition ($p = .033$) and the Individual condition ($p = .020$). No significant difference found between the Pair and the Individual condition.

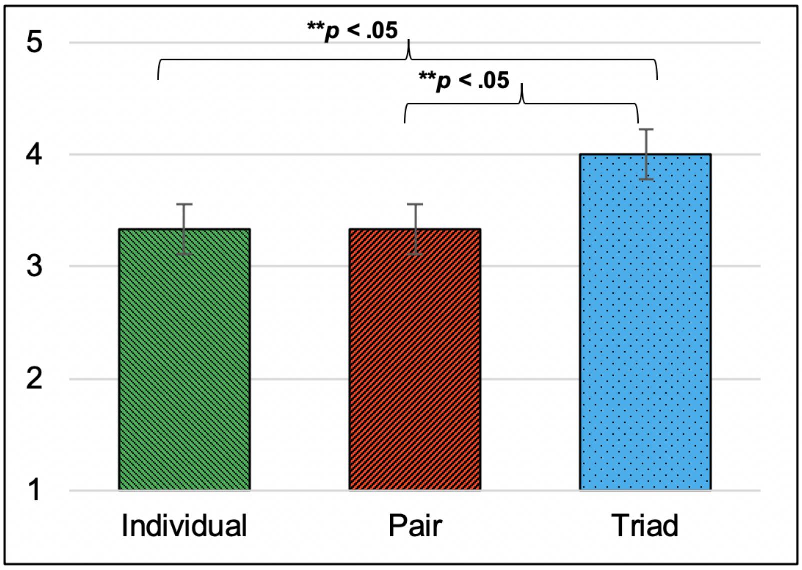


Figure 5.9: Question quality level after refinement feedback

5.2.2 Learning Experiences Questionnaires Result

From the questionnaire data, we conducted a test using a non-parametric Kruskal-Wallis test. Table 5.2 shown the results for each question items. It showed there is a significant difference in the cognitive aspect, specifically in questions item number 9 (*Help Deeper Understanding*) with $p = .046$. Figure 5.11 showed the graph representation of this result. In addition, a Post hoc Mann-Whitney test showed that participants in Triad condition significantly has higher rank than the individual condition (U = 6.000, $p = .041$) and the pair condition (U = 6.000, $p = .046$). No significant difference was found between individual and pair conditions.

Furthermore, we also found that there is a significant effect on the collaborative aspect, which is the question item number 13 (*Prefer Collaborative*) with ($p = .018$). Figure 5.10 showed the graph representation of this result. In addition, a Post hoc Mann-Whitney test showed that the participants in Pair condition (U = 12.000, $p = .016$) and Triad condition (U = 6.000, $p = .018$) prefer collaborative format in question refinement activity than the participants in individual condition.

Table 5.2: Learning experiences questionnaire result

No.	Question Items	Avg. of Individual condition	Avg. of Pair condition	Avg. of Triad condition	p - value
	Emotional Aspect				
1	I felt engaged in this question creation and improvement activity	5.33	5.83	6.17	.166
2	This activity was meaningful	5.83	6	6.17	.775
3	I felt this activity was challenging in a stimulating way	6.67	6.17	6.17	.462
4	This activity was fun learning experiences	6	5.33	6.17	.320
5	This activity was too difficult	3	3.83	3	.316
	Cognitive Aspect				
6	I was able to learn many new things while doing questions improvement	5.5	5.5	5.67	.881
7	I was able to get many new ideas and inspirations while doing questions improvement	5.17	5.67	6	.300
8	This activity helped me to understand the key concepts in the video	5.5	5.33	6.33	.231
9	This activity helped me to have deeper understanding to the concepts introduced in the video	5.33	5	6.33	.046
10	This activity helped me to reflect my understanding to the learning content in the video	5.5	5.33	6.33	.126
11	This activity helped me to memorize the learning content in the video	5.83	5.17	5.33	.686
12	This activity inspired my interests towards the learning content in the video	5.33	4.5	5	.504
	Collaborative Aspect				
13	I prefer doing questions improvement collaboratively with friends	3.83	5.83	6.17	.018
14	I prefer doing questions improvement individually	5	3.167	4.67	.060
15	Having feedbacks is very useful to improve my own questions	6.33	6.17	6.33	.925
	Teaching Material Aspect				
16	The length/duration of the lecture video was good for learning	5.5	5	5.67	.476
17	The length/duration of the lecture video was enough for me making questions	5.33	5.17	6.33	.107
18	The learning content in the video was interesting	5.67	3.67	4.5	.059
19	The learning content in the video increased my curiosity to the topic	5.33	3.83	4	.085
20	The learning content in the video was too difficult for me	3	4.17	2.83	.501
	Workspace Aspect				
21	The workspace (online board) was easy to use for making and refining questions	5.83	5.83	6	.843

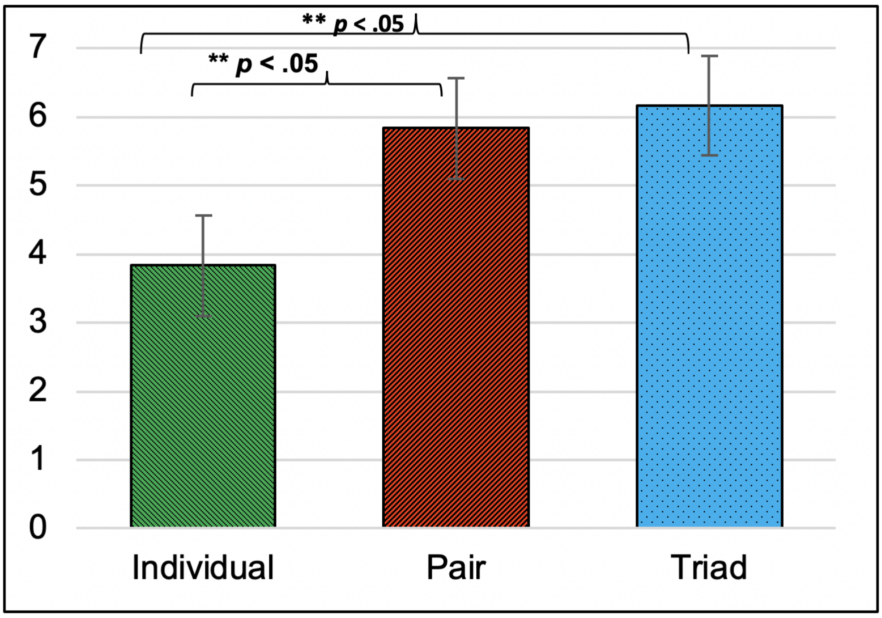


Figure 5.10: Graph representation for the result of question items 9 (prefer collaborative)

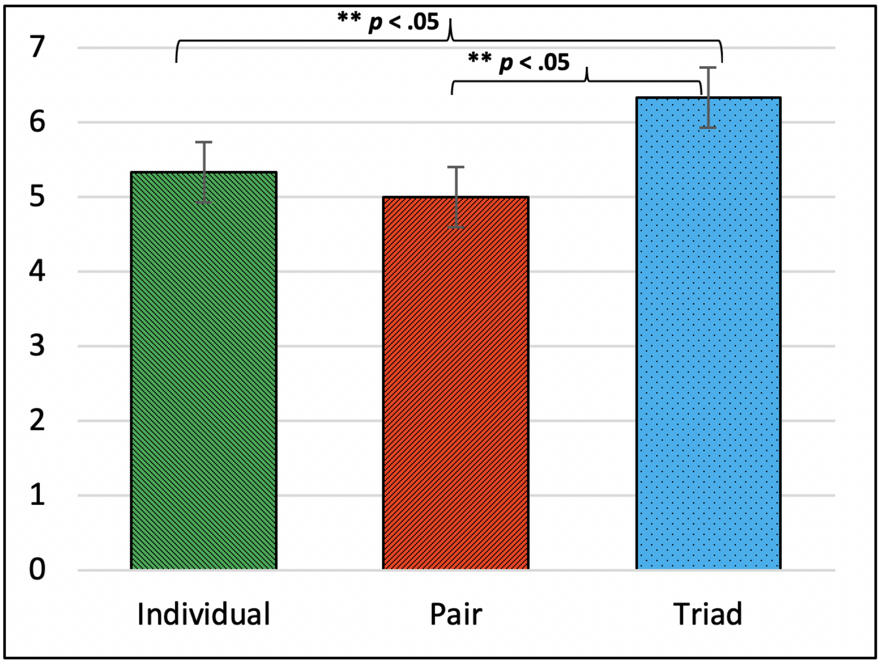


Figure 5.11: Graph representation for the result of question items 13 (deeper understanding)

Chapter 6

Discussion

Results of question quality improvement from the original questions to the refined questions showed that participants in Triad condition can improve the questions to the highest level better than the other conditions. Results from the learning experience questionnaire also showed that the participants in Triad condition felt that they have deeper understanding to the learning content in the video compared to the individual and pair condition. These results suggested that collaborative questions' refinement can increase question quality level as students got feedback in form of refined questions from other students, in which inspired or help them to increase question quality level. This demonstrates a correlation between our study and the claim of Yu [16], who stated that providing feedback was considered as not only helping the question maker to refine their work, but also beneficial to the evaluators because it motivates them to "think critically", and the feedback helped students to find and fix incomplete ideas or misconceptions. Help that can be used to correct someone's questions later.

To gather more insight on the learning by making question refinement activity in this study, we conducted an interview session with participants at the end of the experiment session. This interview data discussed as follows:

Better Understanding. From the interview sessions, we got insight about students learning experience as they stated that they were having better understanding to the video content by doing question generation and refinement activity.

Boosting Knowledge - P5, P6 and P11 stated that this activity assisted them in strengthening their comprehension by relating their prior knowledge (original question) with the new knowledge:

"... I think when you can improve it to other level, you can get a lot deeper answer and it can boost your common knowledge, I think. Not just what it is but also how and why did it happened" (P5).

"... with this kind of method, we will have the opportunity to have a deeper understanding since we are used to make a relationship between one fact concept to another fact" (P6).

"... asking questions and refining the other participant's' questions that can increase our concepts (understanding) about the topic." (P11).

These are associated with previous studies in which students focused their attention and knowledge on making meaning by finding relevant information in the knowledge material and connecting it with their knowledge, which helped improve their understanding [6, 7, 8, 68, 87].

Verifying Knowledge - In addition, on the interview session, P3 claimed that *"... to improve the questions or redefining it expand the meaning from the questions from our self. So, we completely acknowledged what we learned from the video lecture from redefining the question" (P3)*, which indicates how this activity helped him to reconfirm his own knowledge about content of the video. This was also confirmed by P9 *"... (by asking question and refining question) I can find the purpose of a one concept, I mean, I can understand the truly meaning of the concept." (P9)*. This result aligns with Rosenshine's [2] conclusion which also claimed that asking questions help learners to verify their own level of understanding in order to confirm if the content is assimilated.

Based on our observation from the interview results that contain a "better understanding" as one of the benefits. Moreover, we found different reasons regarding students working collaboratively and individually in achieving "better understanding". Students working in groups tend to have reasons that are boosting their knowledge. While students working alone tend to have reasons to have better understanding by verifying their knowledge.

Helps in Remembering Learning Material. While, most of the students saying question generation and refinement helps them in understanding the video, some participants stated that it is not directly helping them. For example, when P17 were asked about whether she got any understanding benefit from this activity she claimed that *"(... I think, not really ... it's not improving like tremendously, (but) like I've thought about this before maybe (helps to remember)" (P17)*. Which indicates it help her to remember the content of the video by generating questions. Also, other students have acknowledged that the question generation and refinement activity could help them to memorize the content, such as P1 and P2:

"... I think it is a kind of way to remember the video content, so it is beneficial for me. By making the questions helps me to remember" (P1).

"... the question refinement like we get to think about what is the content. So, like we memorize a lot more than just watching the videos ..." (P2).

Furthermore, students in pairs condition stated that working collaboratively was the other reason that helps her in remembering the learning material.

"... when I make a question is really a general question and then participant P improved it and then I improve it again, that's makes me remind again about the video (content of the video) ..." (P4).

These are associated with prior works in which students generating high-level questions which require comprehension and application of principles and concepts to new situations is deemed to improve students' learning and retention [7, 8].

Chapter 7

Conclusion

In this study, we examined the process on how students refining questions individually and collaboratively to the improvement level of the refined questions. To investigate these processes, a learning strategy that adapted the three activities from creative processes and student-generated questioning was proposed. A small-scale study comparing students refining questions individually and collaboratively (two students and three students) was conducted using the proposed learning strategy. In this study, the improvement level of the refined question and students' learning experiences were measured.

Based on the result, we found that students refining questions collaboratively in pair have no significant difference in terms of quality level with students refining individually. While when students refining questions collaboratively in triad, they can improve the questions to a better quality.

Moreover, based on the result of students' learning experiences questionnaire, it reported that students in pair and triad condition prefer to refine question collaboratively than individually, and students also claimed to acquire a deeper understanding of the learning material by carrying out question generation and refinement activity.

In addition, students claimed that question generation and refinement activities help them gain deeper understanding (P3, P5, P6, P9, & P11) and help them to memorize the learning material (P1, P2, P4, P17) in the interview sessions. Furthermore, we found different reasons regarding students working collaboratively and individually in achieving "better understanding". Students in collaborative groups tend to gather a more in-depth understanding by boosting/improving their knowledge (e.g. P5). While students working alone tend to have reasons to have better understanding by confirming or verifying their knowledge (e.g. P3).

However, there are some limitations that should be mentioned in this study. First, in this study the number of the participants is too small thus we cannot generalize the current results. For further research, a study with a bigger number of participants should be considered.

Second, the criteria of "unfamiliar with the topic" when recruiting the participants is not specific enough, which might result to having a different "starting level" between participants in terms of comprehension of the learning material. This might affect the result of the generated questions since the deeper the participants' knowledge, the better questions quality they can produced. A suggestion to balance this variable, is by using a pre-test before the experiment to ensure the equality of the comprehension level between participants.

Lastly, the assessment of the question quality in this study is done by the authors themselves without any help from an expert on the learning topic. Even though the taxonomy is clear enough to be used to classify the quality of the questions, there still might be a subjective rating in the assessment. To minimize the subjective variables, an inter-reliability test should ideally be considered to be conducted in the future.

Acknowledgement

Foremost, I would like to express my sincere gratitude to my advisor Prof. Tomoo Inoue for the continuous support of my Master study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Master study.

Besides my advisor, I would like to thank in advance the rest of my thesis committee: Prof. Tomoo Inoue, Prof. Sayan Sarcar , and Prof. Masaki Matsubara, for their encouragement, insightful comments, and hard questions.

My sincere thanks also go to Ari Nugraha, for giving me the chance to join his research group and for all his continuous support and guidance. Also, to his family whose, treated me like a member of their family.

I thank my fellow lab mates in Interaction Lab: Sasaki Kousuke, Harada Tomoyuki, Yuan Zijie, Zhang Jienpieng, and the other lab members who I cannot mentioned here, for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last two years.

I would like to give my thanks to my friends in Tsukuba: Asmae Zaidane, Muhammad Haseeb Khan, and all the Indonesian students in Tsukuba, who always accompanied me when I'm having hard times during these two years.

Also, I would like to thank my friends in Tokyo: Jason Kurniawan and Anom Sulardi for becoming a place for me to share fun and sad moments, and always accompanying and encouraging me every time I have trouble.

Last but not the least, I would like to thank my family: my parents Ahmad Riza Wahono and Noviar, for giving birth to me at the first place and supporting me spiritually throughout my life.

References

[1] Stephen I. Brown and Marion I. Walter. *The Art of Problem Posing*. Psychology Press, 2005.

[2] Barak Rosenshine, Carla Meister, and Saul Chapman. Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, 66(2):181–220, 1996.

[3] Fu-Yun Yu, Yu-Ling Chang, and Hui-Ling Wu. The effects of an online student question-generation strategy on elementary school student English learning. *Research and Practice in Technology Enhanced Learning*, 10(1):24, 2015.

[4] Roslinda Rosli, Mary Margaret Capraro, and Robert M. Capraro. The Effects of Problem Posing on Student Mathematical Learning: A Meta-Analysis. *International Education Studies*, 7(13), 2014.

[5] Bernice Y. L. Wong. Self-Questioning Instructional Research: A Review. *Review of Educational Research*, 55(2):227–268, 1985.

[6] Beth Davey and Susan McBride. Effects of question-generation training on reading comprehension. *Journal of Educational Psychology*, 78(4):256–262, 1986.

[7] Marli E. D. A. Andre and Thomas H. Anderson. The Development and Evaluation of a Self-Questioning Study Technique. *Reading Research Quarterly*, 14(4):605, 1978.

[8] Alison King. Comparison of Self-Questioning, Summarizing, and Notetaking-Review as Strategies for Learning from Lectures. *American Educational Research Journal*, 29(2):303, 1992.

[9] Tracey Papinczak, Raymond Peterson, Awais Saleem Babri, Kym Ward, Vaughan Kippers, and David Wilkinson. Using student-generated questions for student-centered assessment. *Assessment & Evaluation in Higher Education*, 37(4):439–452, 2012.

[10] Albert Logtenberg, Carla Van Boxtel, and Bernadette Van Hout-Wolters. Stimulating situational interest and student questioning through three types of historical introductory texts. *European Journal of Psychology of Education*, 26(2):179–198, 2011.

[11] Alison King. Effects of self-questioning training on college student's comprehension lectures. *Contemporary Educational Psychology*, 14:366–381, 1989.

[12] Chandra J. Foote. Student-Generated Higher Order Questioning as a Study Strategy. *The Journal of Educational Research*, 92(2):107–113, 1998.

[13] Christine Chin and G. Kayalvizhi. Posing problems for open investigations: What questions do pupils ask? *Research in Science and Technological Education*, 20(2):269–287, 2002.

[14] Florian Perteneder, Christian Grossauer, Thomas Seifried, Jagoda Walney, John Brosz, Anthony Tang, Sheelagh Carpendale, and Michael Haller. Idea Playground: When Brainstorming is Not Enough. *Workshop on Designing Collaborative Interactive Spaces (DCIS '12)*, 2012.

[15] Paul Denny, Andrew Luxton-Reilly, and John Hamer. The peerwise system of student contributed assessment questions. 78:69–74, 2008.

[16] Fu Yun Yu. Multiple peer-assessment modes to augment online student question-generation processes. *Computers and Education*, 56(2):484–494, 2011.

[17] Iman Yeckehzaare, Tirdad Barghi, and Paul Resnick. QMaps: Engaging Students in Voluntary Question Generation and Linking. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, New York, NY, USA, 2020. ACM.

[18] Ahmed Elgohary, Denis Peskov, and Jordan Boyd-Graber. Can You Unpack That? Learning to Rewrite Questions-in-Context. pages 5917–5923, 2019.

[19] Ye Liu, Chenwei Zhang, Xiaohui Yan, Yi Chang, and Philip S. Yu. Generative question refinement with deep reinforcement learning in retrieval-based QA system. *International Conference on Information and Knowledge Management, Proceedings*, pages 1643–1652, 2019.

[20] Y. Chen and J. Yin. Apply student-centered learning on computer education. volume 1, pages 465–467, 2011.

[21] Ali Alharbi, Frans Henskens, and Michael Hannaford. Student-Centered Learning Objects to Support the Self-Regulated Learning of Computer Science. *Creative Education*, 03(06):773–783, 2012.

[22] Jane E. Brindley, Christine Walti, and Lisa M. Blaschke. Creating effective collaborative learning groups in an online environment. *International Review of Research in Open and Distance Learning*, 10(3):1–18, 2009.

[23] Jan Hawkins and Roy D. Pea. Tools for bridging the cultures of everyday and scientific thinking. *Journal of Research in Science Teaching*, 24(4):291–307, 1987.

[24] J. L. Lemke. *Talking Science: Language, Learning, and Values*. Ablex Publishing Corporation, Norwood, 1990.

[25] Yuuki Hirai, Atsuo Hazeyama, and Tomoo Inoue. Assessment of learning in concerto iii: a collaborative learning support system based on question-posing. *Proceedings of Computers and Advanced Technology in Education. Calgary, Canada: ACTA Press*, pages 36–43, 2009.

[26] Sam Saarinen, Shriram Krishnamurthi, Kathi Fisler, and Preston Tunnell Wilson. Harnessing the Wisdom of the Classes. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 606–612, New York, New York, USA, 2019. ACM Press.

[27] Arthur C. Graesser and Natalie K. Person. Question Asking During Tutoring. *American Educational Research Journal*, 31(1):104–137, 1994.

[28] Linda Elder and Richard Paul. Universal intellectual standards, 1999.

[29] Alison King. Enhancing Peer Interaction and Learning in the Classroom Through Reciprocal Questioning. *American Educational Research Journal*, 27(4):664–687, 1990.

[30] R. Fisher. *Teaching Children to Think*. Prentice Hall, 1990.

[31] M Browne and S. M. Keeley. *Asking the right questions: A guide to critical thinking*. Pearson/Prentice Hall, 5th edition, 1998.

[32] Helena Pedrosa de Jesus, José J.C. Teixeiera-Dias, and Mike Watts. Questions of chemistry. *International Journal of Science Education*, 25(8):1015–1034, 2003.

[33] M. A. Corley and W. C. Rauscher. TEAL Fact Sheet No. 12: Deeper Learning through Questioning. *Teaching Excellence in Adult Literacy*, 12:1–5, 2013.

[34] Benjamin S. Bloom. *Taxonomy of educational objectives: the classification of educational goals*. Longman Group, 1956.

[35] L. Anderson, D. Krathwohl, P. Airasian, and B. Samuel. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives Longman, New York 2001*. 2001.

[36] J. Gallagher and M. Aschner. A Preliminary Report on Analysis of Classroom Interaction. *Merrill-Palmer Quarterly of Behavior and Development*, 9:183–194, 1963.

[37] Peter Mosenthal. Understanding the strategies of document literacy and their conditions of use. *Journal of Educational Psychology*, 88:314–332, 1996.

[38] Ana Taboada and John T. Guthrie. Growth of cognitive strategies for reading comprehension. *Motivating Reading Comprehension: Concept-Oriented Reading Instruction*, pages 232–259, 2004.

[39] David Wood, Jerome S Bruner, and Gail Ross. The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, 17(2):89–100, 1976.

[40] Shari Jackson, Joseph Krajcik, and Elliot Soloway. The design of guide learner-adaptive scaffolding in interactive learning environments. pages 187–194, 1998.

[41] Michael Hannafin and Susan Land. Technology and student-centered learning in higher education: Issues and practices. *Journal of Computing in Higher Education*, 12:3–30, 2012.

[42] Xun Ge and Susan Land. A conceptual framework for scaffolding ill-structured problem-solving processes using question prompts and peer interactions. *Educational Technology Research and Development*, 52:5–22, 2004.

[43] Carl Bereiter and Marlene Scardamalia. Cognitive coping strategies and the problem of inert knowledge. *Thinking and Learning Skills: Current Research and Open Questions*, 2:65–80, 1985.

[44] Alison King. A strategy for enhancing peer interaction and learning during teacher training sessions. *Teacher Education Quarterly*, 1991.

[45] Cheryl Weiner. The effect of training in questioning and student question generation on reading achievement. 1978.

[46] M.J. Dreher and L.B. Gambrell. Teaching children to use a self-questioning strategy for studying expository prose. *Reading Improvement*, 22:2–7, 1985.

[47] Bernice Wong and Wayne Jones. Increasing metacomprehension in learning disabled and normally achieving students through self-questioning training. *Learning Disability Quarterly*, 5:228–240, 1982.

[48] P. Ritchie. The effects of instruction in main idea and question generation. *Reading Canada Lecture*, 3:139–146, 1985.

[49] Rosemary Lonberger. The effects of training in a self-generated learning strategy on the prose-processing abilities of fourth and sixth graders. 1989.

[50] T. Raphael and P. Pearson. Increasing students' awareness of sources of information for answering questions. *American Educational Research Journal - AMER EDUC RES J*, 22, 1985.

[51] M. Dermody. Metacognitive strategies for development of reading comprehension for younger children. 1988.

[52] G. Labercane and J. Battle. Cognitive processing strategies, self-esteem, and reading comprehension of learning disabled students. *B. C. Journal of Special Education*, 11, 1987.

[53] N. J. Smith. *The effects of training teachers to teach students at different reading ability levels to formulate three types of questions on reading comprehension and question generation ability. Unpublished*. PhD thesis, University of Georgia, 1977.

[54] R.Y. Nolte and H. Singer. Active comprehension: Teaching a process of reading comprehension and its effects on reading achievement. *The Reading Teacher*, 39:24–31, 1985.

[55] E. Short and E. Ryan. Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attribution training. *Journal of Educational Psychology*, 76:225–235, 1984.

[56] Denis Duret, Rob Christley, Paul Denny, and Avril Senior. Collaborative learning with peerwise. *Research in Learning Technology*, 26, 2018.

[57] Stephen Snow, Adriana Wilde, Paul Denny, and m.c Schraefel. A discursive question: Supporting student-authored multiple choice questions through peer-learning software in non-stemm disciplines. *British Journal of Educational Technology*, 50, 2018.

[58] Chun-Ping Wu, Yu-Chi Chen, and Shu-ling Wu. Development of a collaborative learning space for student generated-question strategy. pages 102–105, 2018.

[59] D. Hancock, N. Hare, P. Denny, and G. Denyer. Improving large class performance and engagement through student-generated question banks: Student-generated question banks. *Biochemistry and Molecular Biology Education*, 46, 2018.

[60] S. Yang, H. Zhewei, and E. F. Gehringer. Who Took Peer Review Seriously: Another Perspective on Student-Generated Quizzes. *EDM*, 2016.

[61] Paul Denny. Generating practice questions as a preparation strategy for introductory programming exams. *SIGCSE 2015 - Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, pages 278–283, 2015.

[62] D Biggins, E Crowley, E Bolat, M Dupac, and H Dogan. Using Peerwise to Improve Engagement and Learning. page 16, UK, 2015. The European Conference on Education 2015.

[63] P. Denny, E. Tempero, D. Garbett, and A. Petersen. Examining a student-generated question activity using random topic assignment. pages 146–151, 2017.

[64] Simon Bates, Ross Galloway, and Karon McBride. Student-generated content: Using peerwise to enhance engagement and outcomes in introductory physics courses. *AIP Conference Proceedings*, 1413:123–126, 2012.

[65] S. Bottomley and P. Denny. A participatory learning approach to biochemistry using student authored and evaluated multiple-choice questions. *Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology*, 39:352–61, 2011.

[66] A. Sharp and A. Sutherland. Learning Gains... "My (ARS)" The impact of student empowerment using Audience Response Systems Technology on Knowledge Construction, Student Engagement and Assessment. In *Student Engagement and Assessment REAP International Online Conference on Assessment Design for Learner Responsibility,*, pages 29–31, 2007.

[67] D. Nicol. E-assessment by design: Using multiple-choice tests to good effect. *Journal of Further and Higher Education*, 31, 2007.

[68] Heather Mcqueen, C. Shields, D. Finnegan, J. Higham, and M. Simmen. Peerwise provides significant academic benefits to biological science students across diverse learning tasks, but with minimal instructor intervention. *Biochemistry and Molecular Biology Education*, 42, 2014.

[69] A. Luxton-Reilly, D. Bertinshaw, P. Denny, B. Plimmer, and R. Sheehan. The impact of question generation activities on performance. *SIGCSE'12 - Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*, 2012.

[70] J. Hardy, S. Bates, M. Casey, K. Galloway, R. Galloway, A. Kay, P. Kirsop, and H. Mcqueen. Student-generated content: Enhancing learning through sharing multiple-choice questions. *International Journal of Science Education*, 36, 2014.

[71] Lawrence Frase and Barry Schwartz. Effect of question production and answering on prose recall. *Journal of Educational Psychology*, 67:628–635, 1975.

[72] Stephen Draper. Catalytic assessment: Understanding how mcqs and evs can foster deep learning. *British Journal of Educational Technology*, 40:285 – 293, 2009.

[73] M. Fellenz. Using assessment to support higher level learning: the multiple choice item development assignment. *Assessment Evaluation in Higher Education*, 29, 2005.

[74] Paul Denny. The effect of virtual achievements on student engagement. pages 763–772, 2013.

[75] S. Miller and C. Brownell. Peers, persuasion and piaget: Dyadic interaction between conservers and nonconservers. *Child Development*, 46:992, 1975.

[76] Igor Podgorny, Matthew Cannon, Chris Gielow, and Todd Goodyear. Real time detection and intervention of poorly phrased questions. pages 2205–2210, 2015.

[77] Bill Buxton. *Sketching user experiences: getting the design right and the right design*. Morgan kaufmann, 2007.

[78] Alex Osborn. *Applied imagination-principles and procedures of creative writing*. Read Books Ltd, 2012.

[79] Ester Aflalo. Students generating questions as a way of learning. *Active Learning in Higher Education*, 2018.

[80] John Guthrie and Nicole Scafiddi. Reading comprehension for information text: Theoretical meanings, developmental patterns, and benchmarks for instruction. *Motivating Reading Comprehension: Concept-Oriented Reading Instruction*, 2004.

[81] John T. Guthrie, A. Laurel W. Hoa, Allan Wigfield, Stephen M. Tonks, Nicole M. Humenick, and Erin Littles. Reading motivation and reading comprehension growth in the later elementary years. *Contemporary Educational Psychology*, 32(3):282–313, 2007.

[82] Dongsong Zhang, Lina Zhou, Robert Briggs, and Jay Nunamaker. Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information Management*, 43:15–27, 2006.

[83] Meg Colasante. Using video annotation to reflect on and evaluate physical education pre-service teaching practice. *Australasian Journal of Educational Technology*, 27, 2011.

[84] Franck Ganier and Philine de Vries. Are instructions in video format always better than photographs when learning manual techniques? The case of learning how to do sutures. *Learning and Instruction*, 44:87–96, 2016.

[85] Miriam Gamoran Sherin and Elizabeth A. Van Es. Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60(1):20–37, 2009.

[86] Mohamed Amine Chatti, Momchil Marinov, Oleksandr Sabov, Ridho Laksono, Zuhra Sofyan, Ahmed Mohamed Fahmy Yousef, and Ulrik Schroeder. Video annotation and analytics in CourseMapper. *Smart Learning Environments*, 3(1), 2016.

[87] Mehdi Alaimi, Edith Law, Kevin Daniel Pantasdo, Pierre-Yves Oudeyer, and Hélène Sauzeon. Pedagogical Agents for Fostering Question-Asking Skills in Children. pages 1–13, 2020.

[88] Yushun Li, Meiqin Wu, Jinrui Dai, and Shengri Chen. Blended Learning. New Challenges and Innovative Practices. 10309:84–96, 2017.

[89] John Keller and Katsuaki Suzuki. Learner motivation and e-learning design: A multinationally validated process. *Learning, Media and Technology*, 29:229–239, 2004.

[90] Sunyoung Joung and John Keller. The effects of high-structure cooperative versus low-structure collaborative design of decision change, critical thinking, and interaction pattern during online debates. *Association for Educational Communications and Technology*, 2004.

[91] Philip J. Guo, Juho Kim, and Rob Rubin. How video production affects student engagement. In *Proceedings of the first ACM conference on Learning @ scale conference - L@S '14*, pages 41–50, New York, New York, USA, 2014. ACM Press.

[92] Tingting Yang and Heng Luo. Learning Experience Questionnaire for Multimedia Case-Based Instruction: Development, Validation and Implications for Design. In Simon K S Cheung, Lam-for Kwok, Will W K Ma, Lap-Kei Lee, and Harrison Yang, editors, *Blended Learning. New Challenges and Innovative Practices*, pages 400–411, Cham, 2017. Springer International Publishing.

[93] Kth Royal Institut of Technology. Learning experience questionnaire (leq), 2020.