Self-Explanation and Self-Questioning Prompts in Online Medical Health Learning

by

Donya Samadi

B.Sc., University of British Columbia, 2013

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

in the
Educational Psychology Program
Faculty of Education

SIMON FRASER UNIVERSITY
Spring 2017

Approval

Name:	**Donya Samadi**
Degree:	**Master of Arts**
Title:	**Self-Explanation and Self-Questioning Prompts in Online Medical Health Learning**
Examining Committee:	**Chair:** **Peter Liljedahl** Associate Professor
	Philip Winne Senior Supervisor Professor
	John Nesbit Supervisor Professor
	Margaret MacDonald Internal/External Examiner Associate Professor
Date Defended/Approved:	April 13, 2017

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Abstract

Online instruction in medical education is beneficial due to moves toward competency-based curricula, continuing education, serving professionals in remote locations, and knowledge updates as research advances. Those who study content online may require support to use effective methods that transform passive, less-engaged learning into active comprehension and purposeful application. This study compared two learning tactics: self-questioning and self-explanation that have not been compared in prior research.

Health professionals and students across Canada studied a chapter in the *Canadian Fundamentals of Fetal Health Surveillance (FHS) Self-Learning Online Manual*, presented on an online learning management system. Participants used nStudy learning software to open note templates and type in either self-explanations or choose one among several question stems then fill in blank space(s) to create a question.

Participants who created self-explanations performed better on the achievement posttest than those who generated self-questions. Further analyses disaggregated posttest items into intentional learning (relating to information in the text about which participants were prompted to generate an annotation) and incidental learning (relating to information in the text not directly prompted for annotation). Within the self-explanation condition, there was no statistically detectable difference in recall on intentional (prompted) content compared to incidental (non-prompted) content. In the self-questioning condition, incidental content was recalled similarly to the self-explanation group. However, there was a marked and statistically detectable decrease in recall of content about which participants were prompted to generate self-questions. Possible reasons for this effect based on past research and participant comments are discussed along with limitations of the study and opportunities for further research.

Keywords: self-questioning; self-explanation; medical education; nursing education; prompts; recall

To my parents for all their love, dedication, and support.

Acknowledgements

I had the privilege of learning from a truly gifted teacher, my supervisor Dr. Phil Winne. Thank you, Dr. Winne, for your consistent encouragement, advice, and comments. I learned a lot from you throughout these years and am deeply grateful for the opportunity to work with such a wonderful educator and researcher. Thank you to Dr. John Nesbit for your thoughtful feedback and valuable knowledge, your guidance was greatly appreciated. I would like to thank Janet Walker for supporting my use of the FHS Manual for this study, Sue Paine for all her help in gathering participant registrations, and everyone at CESEI who gave me an introduction to the area of online medical learning.

I thank my parents, Qamar and Gite Samadi, without whom this would not be possible – they have supported and encouraged me throughout my entire educational journey and continue to do so. Thank you to my brothers, Omron and Omar, for your enthusiasm, support, and patience. Thank you, Ali, for your love, inspiration, and humor. Thanks to my lovely cousins, friends, and all my family for your belief in me and encouragement to continually pursue more challenging goals. I would also like to thank my nStudy lab members, it was a joy to work with you all and I am glad to have met each of you – Zahia, I appreciate all your advice and kindness. I have been blessed to be surrounded by a great support network, both at home and in the university, and this positively impacted my journey to successfully complete my Master's research as well as encouraging my pursuit of future endeavors.

Table of Contents

List of Tables

List of Figures

List of Acronyms

FHS	Fetal Health Surveillance
LMS	Learning Management System
SQ	Self-questioning
SE	Self-explanation
BN	Blank note

Chapter 1.

Introduction

Online learning is becoming more prevalent for learners in all disciplines. Online instructional materials allow students to control the time and pace of learning, the sequence in which they access content (via hypermedia links and search), and the mode of annotating material to be learned, making it highly individualized and accessible. Online learning in medical education has become especially popular in recent years due to moves toward competency-based curricula and continuing lifelong medical education, a need for distance learning to serve those in remote locations, and a requirement that knowledge be updated constantly due to advances in research (Zehry, Halder, & Theodosiou, 2011). A survey of over 1,000 medical students and licenced doctors disclosed they were interested in using web technology in medical education, but training in how to use the technology and online resources was necessary to comfortably incorporate online learning in their medical studies (Sandars & Schroter, 2007). A review of web-based medical education studies revealed that online instruction was just as beneficial to learners as traditional means of instruction (e.g., textbooks, lectures) and, in some cases, it was more efficient (Chumley-Jones, Dobbie, & Alford, 2002).

Self-directed learning occurs when individuals determine what they need to learn, and decide how they will achieve their learning goals (Knowles, 1975). A focused review of self-directed learning (SDL) in health professions education found that, compared to traditional instruction, SDL was associated with better knowledge-based learning outcomes (Murad, Coto-Yglesias, Varkey, Prokop, & Murad, 2010). When self-directed learning is an active process, students are more motivated and engaged with the content (Ruiz, Mintzer, & Leipzig, 2006) because they are interacting with instructional multimedia in a flexible way according to their own preferences. However, for self-directed learning to be as efficient as possible, learners may need to be provided with guidance so as not to miss important information. Aside from potential benefits of learner control while studying, presenting instruction online allows collecting fine-grained data about a learner's behaviour and outcomes. These data can be used by instructors to design and implement interventions to enhance both the process and product of learning.

The instructional content used in this study is the Canadian Fundamentals of Fetal Health Surveillance Self-Learning Online Manual. This evidence-based, expert-reviewed, fetal health surveillance (FHS) manual is designed as pre-reading for health professionals and students (including medical students, nursing students, and midwifery students) who must update their FHS knowledge every two years before attending an in-person workshop to consolidate and further their learning. The manual was originally a hard-copy textbook mailed to all practitioners who required it. However, in 2011, a survey indicated 90% of FHS instructors were more interested in an online version, with interactive modules (Perinatal Services BC, 2014). Therefore, the content was transferred into online modules on a learning management system, with some modules including images, interactive animations, and audio. The aim of transferring this manual to a web-based format on an LMS was to improve accessibility, efficiency, be cost-effective, and allow for more frequent and easier updates of the content. The stimulus for this thesis research was to explore how learners use the online FHS manual, and investigate opportunities through which learning can be enhanced, not just for these professionals but for all online learners.

When passively learning, students take in knowledge without engaging to think about how this knowledge fits into their pre-existing schemas (Graffam, 2007). In contrast, active learning involves students purposefully operating on information. Strategies used in active learning include summarizing information, integrating it with prior knowledge, and asking questions (Graffam, 2007; Mayer, 2010). In an online learning environment, the student must make decisions about how to learn, how much to learn, how much time to spend, and whether material has been understood (Azevedo & Cromley, 2004). Previous studies showed students often use ineffective strategies in such situations and require some support (Azevedo & Cromley, 2004). It is important for instructors to maximize how well students make sense of information while reducing extraneous cognitive processing (Mayer, 2010). One way to do this is introducing prompts to guide learners as they read the online content. Providing instructional support via prompts is valuable, especially in science topics which contain nonlinear, multilayered complex relations (Azevedo, 2005). Simply having students take notes as they learn online is not as beneficial as providing them with specific prompted notes because when learners are unguided they tend to write out shallow reproductions of the content and learn less (Trevors, Duffy, & Azevedo, 2014). Prompts that ask students to produce a specific type of annotation might promote learning compared to non-prompted online learning. My thesis research explored the cognitive benefits of providing question stems (for self-questioning) and prompting students to explain a concept to themselves (i.e., self-explanation).

The implications of this study will inform decisions about types of prompts that may enhance students' learning. Results could direct future research into adaptive instruction, using the most effective prompt types.

Chapter 2.

Review of the Literature

2.1. Self-Questioning

Teaching students to create questions while studying new material increases comprehension (Rosenshine et al., 1996). In a review of self-questioning research, Wong (1985) found the technique of students generating their own questions as they read a text improved student comprehension.

Cano García et al. (2014) explored the effect of learner-generated questions when studying a science topic. Some studies show that generating questions during learning aids in active processing of information, while other studies use learner generated questions to promote monitoring comprehension (Cano Garcia et al., 2014). Results indicated students in the prompted self-questioning group had greater gains relative to pretest scores, increased metacognitive knowledge and more self-regulation of strategies than students not provided prompts to self question (Cano Garcia et al., 2014).

The physical act of typing out answers to questions affects a student's learning. A study giving prompts to learners compared knowledge acquisition and transfer between a group who merely received question prompts to think about and answer to themselves, versus a group who typed their answers to the question prompts (Papadopoulos, Demetriadis, Stamelos, & Tsoukalas, 2011). Students told to think about answers to the prompts said they tended to either skip the prompts, or reported reading them but had lower performance in posttest domain knowledge and in transfer of knowledge to novel problems than students who typed answers to the prompts (Papadopoulos et al., 2011). Thus, incorporating answering by the student appears to actively engage them and improve learning.

When comparing the strategies of summarizing content after a lecture to generating questions (and answering them), King (1992) found students who wrote summaries recalled more on an immediate posttest than those generating self-questions.

However, when tested again one week later, those who had generated and answered their own questions had higher retention than those who had produced summaries, suggesting the strategy of self-questioning is more effective than summarizing for long term recall. Both strategies resulted in higher performance than the control group who only reviewed their lecture notes after class.

Duell (1977) compared two groups of college students who were given learning objectives. One group was instructed to study a text based on the objectives and the other was instructed to write questions related to the objectives. Learners who created questions performed better on both a recognition test and an application test (Duell, 1977).

2.2. Theoretical Perspectives of Self-Questioning

Three theoretical perspectives underlie self-questioning research according to Wong (1985). One is active processing, which assumes that students who generate questions to guide their cognitive activity are more aware and attentive. Those following this theoretical perspective believe students should form questions for themselves as they study to ensure that they are actively learning. Criticisms of this perspective posit the type of self-questions created must match the need of the student; for example, students who need to focus on key words should be prompted by teachers/experimenters to generate questions focusing on key words.

From the metacognitive perspective, self-questioning is beneficial not when students are just told to create questions for themselves as they read but when students are taught to monitor their understanding and determine for themselves when it would be appropriate to generate a question about particular content (Wong, 1985). In open-ended online learning environments, where students must adequately use metacognitive skills to learn (Azevedo, 2005), self-questioning can be categorized as a metacognitive learning strategy because students think must metacognitively monitor what they know and what they don’t know to form questions (King, 1992). Self-questioning can also be cognitive, in that when students attempt to answer a question they have generated, they are testing themselves on the content knowledge (King, 1992).

From the third perspective, schema theory, self-questioning activates students' relevant prior knowledge which improves text comprehension (Wong, 1985). King (1994) found questions which prompt a learner to access prior knowledge related to the content being studied are more effective than questions which ask solely about the ideas presented. From a constructivist viewpoint, individuals make meaning for new information by drawing on their previous knowledge to build inferences, elaborations, and create relationships between the new information and knowledge (King, 1994). Students prompted to construct high level integrating questions gave more complex integrated statements while students given prompts to generate basic factual questions had low knowledge construction by simply restating information (King, 1994). When the question prompts were removed, learners still formed more integration questions than the control group. This shows question strategies can be learned although the students also asked more basic factual questions and engaged in simpler knowledge construction than questions were explicitly prompted (King, 1994).

The self-questioning prompts used in the current research adhere closest to the active processing theory and schema theory (Janssen, 2002). The link to schema theory is because learners are instructed to think about "how concepts in the text link to information you already know." The question stems also prompt relations to prior knowledge. Compared to questions that encourage connections between ideas in a lesson, it seems that questions which prompt accessing prior knowledge lead to greater comprehension (King, 1994).

The self-questioning in my study also relates to active processing theory because learners were prompted at regular intervals to generate questions, sustaining active engagement with the text. Generating self-questions seems to ensure that learners are actively involved in the material they are reading (Burnett & Berg, 1988).

The participants in my research were prompted to create cognitive rather than metacognitive questions for themselves. Students may implicitly be detecting their comprehension and what kind of a question they should be asking; however, the nature of the study requires more cognitive type questions to be able to compare to the self-explanation and no-prompt conditions.

2.3. Instruction in Self-Questioning

Students need instruction about how to generate self-questions. Andre and Anderson (1978) found students who were given training generated more and better questions compared to an un-trained group. However, both questioning groups performed similarly on a posttest and outperformed groups who only reread the text (Andre & Anderson, 1978).

It is unclear how much instruction is optimal. Rosenshine et al. (1996) compared a wide range of training durations (from 2 hours to 12 hours). Self-questioning success was not clearly associated to time allotted to training or the number of training sessions provided. However, other studies show more training sessions in questioning result in higher comprehension (Huang, 1992; Westera & Moore, 1995).

In the current study, training was basic: participants were provided a short paragraph about self-questioning before initiating study. Due to the nature of the content presentation, solely online, it makes sense to have the self-questioning instruction presented online as well.

Aside from the type of prompts given to students in self-questioning instruction and the duration of the training, other variations in instructional approaches were equally effective in improving comprehension and recall, e.g.: reciprocal questioning vs. conventional instruction, receiving assistance from teachers vs. peers, and cooperative learning vs. cross-age tutoring (Janssen, 2002).

2.4. Self-Questioning Prompts

Different types of self-questioning prompts have been given to students in a variety of research studies. Generic question stems and signal words such as "Why" or

“How” in prompts have been found to be most effective when compared to other types of prompting interventions such as focusing on question types (e.g. inference questions vs. factual questions) or questions that prompt students to identify main ideas (Rosenshine et al., 1996). King (1994) found that experience-based questioning was more beneficial to students than content-based questioning.

In some studies, students were asked to answer questions they generated (King, 1992; King, 1994). In other studies, students were instructed to generate self-questions but not answer the questions they had created (Mostow & Chen, 2009; Andre & Anderson, 1978; Duell, 1977). In this study, I decided to focus on question generation, rather than response. Students are provided with question stems along with blank spaces to completely form the question. However, there is no space for a question response, and students can choose to create a reply to themselves as they like. The reason the question response was not included in this study was because not responding to self-generated questions provides the greatest contrast to the other condition in this study, self-explanation, which requires a generated explanation in response to a prompt.

2.5. Self-Explanation

In self-explanations, learners describe why a concept or idea is important and how it relates to their prior knowledge (Larsen et al., 2013). Another version of self-explanation has been studied when content is a step-by-step procedure and learners must explain why a subsequent step follows a prior step (Siegler, 2002). Learners infer causal or conceptual connections when generating self-explanations (Bisra, Liu, Salimi, Nesbit, & Winne, 2016). Berry (1983) compared a group of participants who generated self-explanations during a learning task to a group who self-explained after the task had been completed. Performance was greater for those who formed self-explanations concurrently with the task (Berry, 1983). When students are told to generate self-explanations, they may not follow directions and consequently produce relatively few self-explanations (Hausmann & Chi, 2002). However, when they are prompted, the number of self-explanations increases and this correlates with improvements in learning (Hausmann & Chi, 2002). Even when feedback is not given to students who generate prompted self-explanations, learning improves (Schworm & Renkl, 2006). A meta-analysis by Bisra et al. (2016) found self-explanation prompts to be an effective intervention to improve learning outcomes, with an overall weight mean effect size using a random effects model of $g = 0.54$. Comparing the placement of prompts, those studies in which self-explanation prompts were presented throughout the learning content had a larger, statistically detectable effect size than prompts given before or after the learning session (Bisra et al., 2016).

Wong, Lawson, & Keeves (2002) investigated how the presence of self-explanation prompts affected high school students' problem solving in math. Prompts for self-explanations were questions in each section, such as "What does the statement mean?" and "How does this new piece of information help me to solve the sample problem?" Students in the self-explanation condition had higher posttest scores than peers in the control condition despite similar pretest scores. The greatest difference in achievement was found on transfer questions where students had to generalize knowledge they had studied to construct new figures and use theorems not present in the study booklet (Wong et al., 2002). Thus, the learning effect of self-explanation applies not just to general knowledge recall but also to transfer of knowledge.

Trabasso and Magliano (1996) looked at think aloud protocols and found that, compared to generating predictions or making associations while reading text, students were more likely to generate explanations, concluding that comprehension of text is mostly explanation based.

Self-explanation is seen to be an effective learning strategy for a few reasons (Chi, Leeuw, Chiu, & LaVancher, 1994). Generating self-explanations is an active, constructive process in which the learner creates knowledge. Also, during self-explanation, a learner is integrating new information with pre-existing knowledge, as opposed to other constructive activities such as summarizing that may not assemble new information with prior knowledge (Chi et al., 1994). Wittrock's model of generative learning describes that when students interact with content and use their previous experience and knowledge to construct meaning, they are better able to understand and remember the content (King, 1992). Reformulating and adding to information builds cognitive structures which facilitate understanding and memory (King, 1992). Finally, self-explanation can occur many times over the course of learning, and learners have the opportunity to go back to previous generations of explanations and revise any incorrect assumptions as they move through a text and gain more knowledge (Chi et al., 1994).

Having students generate self-explanations results in better recall and application than studying without self-explanations (Larsen, Butler, & Roediger III, 2013). In Larsen et al.'s study, the self-explanations were generated a week after the teaching session, and retention was measured six months after the studying sessions (Larsent et al., 2013). For those learning online, it may not be practical to learn material, and then return to the material a week later to generate self-explanations. It would be beneficial to explore immediate generation of self-explanations and investigate whether this results in improved learning compared to a group who does not generate self-explanations.

2.6. Comparing Self-Explanation and Self-Questioning

Larsen et al. (2013) compared generation of self-explanations to repeated testing. Repeated testing resulted in better retention six months later than self-explanations. However, students are often pressed for time when studying for several courses, and repeated testing of material, although beneficial, might not be realistic in such settings because of the additional time and effort investment required to create many test questions. Rather than providing the student with test questions, asking them to construct their own questions could be a way for students to learn how to self-question, and connect the material with their prior knowledge and experience. Prompting a student to come up with their own explanations and questions while reading and studying allows the learner to practice using such strategies, and after continuous use the student may decide to carry over such effective learning tactics to their other courses and study sessions.

Another point to consider is that creating questions as a learning strategy is effective because it allows the learner to practice retrieval of information (if they decide to answer the questions they have generated), thus enhancing learning and retention both short-term and long-term (Roediger & Karpicke, 2006). Similarly, self-questioning prompts invite the learner to think back on what they have read and use the information to form questions. This activity may enhance learning by activating retrieval practice, as a learner decides what information is most pertinent to form a question about.

A review of the research did not reveal any studies specifically comparing self-explanation to self-questioning. Therefore, this study investigates whether each of self-questioning and self-explanation improve learning beyond the control group who do not receive specific prompts, and whether there is a difference in test achievement depending on the type of prompt received. Other factors to be taken into consideration include the role of prior knowledge and professional background.

Research Question:

Does a participant's posttest achievement depend on type of prompts received: prompts to self-explain or prompts to generate questions?

Chapter 3.

Method

3.1. Participants

Medical professionals involved in perinatal care of infants in the weeks shortly before and shortly after birth are required to be knowledgeable in fetal health surveillance. Practitioners serving pregnant women must first receive certification through an in-person one-day Fundamentals of Fetal Health Surveillance (FHS) Provider workshop. In advance of this workshop, students and professionals are required to study *The Fundamentals of Fetal Health Surveillance: A Self-Learning Manual.*

Participants in this study were recruited through a contact at Perinatal Health Services in British Columbia. Instructors of the FHS workshop from hospitals and educational institutions across Canada request access to the online FHS Manual from Perinatal Health Services BC, and these instructors as well as their students were sent a consent form describing participation in this study. Participants thus represent a variety of educational backgrounds and experience, although all had in common that they were required to read the FHS online manual to further their education and work. Participants were 45 health care professionals or students, 44 of them female and 1 male. Of the 45 participants, 19 were registered nurses, 5 midwives, 3 medical doctors, 5 nursing students, 9 student midwives, and 4 medical students. Appendix A contains a breakdown of the profession of participants by condition.

3.2. Materials

The instructional content in this study is the *Canadian Fundamentals of Fetal Health Surveillance Self-Learning Online Manual.* This evidence-based, expert-reviewed, fetal health surveillance (FHS) manual is pre-reading for medical, nursing, and midwifery students, and health professionals who must update their FHS knowledge every two years before attending an in-person workshop to consolidate and further their learning. The manual is password-protected and available to students from the *Centre of Excellence for Simulation Education & Innovation* (CESEI) learning management system. For this study, students read Chapter 5 (Intermittent Auscultation (IA) Of The Fetal Heart Rate (FHR)) of the online manual while using nStudy software.

nStudy is an online learning software system designed by Dr. Phil Winne and Dr. John Nesbit at Simon Fraser University. Using nStudy, learners can study online content using a variety of features, including: highlighting text, annotating and tagging content, organizing artifacts, and collaborating with others. In this study, participants are given instructions about using nStudy, specifically selecting text and creating notes (see Appendix B for example of nStudy tutorial given to participants). Users logged into nStudy have their studying behaviour on all CESEI LMS pages recorded and time-stamped, including tabs switched, and text operated on (e.g. selected/note template opened/text entered into note).

3.3. Measures

A questionnaire gathered demographic information about participants' age, gender, prior education, relevant work experience, and exposure to obstetrical knowledge.

A test of prior knowledge showed 9 key concepts related to the Chapter 5 content. Participants checked "the ones that you think you could describe well to a classmate." A space was provided for each participant to "give a brief description of the concepts you think you know about." (See Appendix C for the Demographic Questionnaire and Prior Knowledge test.) This format, rather than a blank open-ended question, was intended to limit the scope of participant responses to topics covered in Chapter 5 and allow comparison of prior knowledge to posttest items. Differential priming was not a concern as participants in all three conditions were exposed to the same prior knowledge test.

An interpolated task was administered following study but before testing. This questionnaire asked about familiarity with online learning environments and general technology skills, etc. (Appendix E).

A posttest of achievement was developed for Chapter 5, the content participants studied. It contained 18 prompted recall questions requiring participants to give short answers or point form lists. The order of posttest questions was randomized to avoid primacy and recency effects. A scoring rubric for responses to the recall questions (Appendix F) was developed to identify information in Chapter 5 necessary to answer each question. Eight questions targeted information in the text for which participants were prompted to generate a note. These items will be referred to as items targeting intentional learning because material needed to answer them was directly referenced by the prompts. Ten of the questions asked about content which was not prompted. Content to answer these items is identified as incidental learning as it was not directly prompted. Comparing recall of prompted material to non-prompted material allowed exploring the efficacy of prompts.

All four measures were hosted on a version of FluidSurveys (online survey tool) for Simon Fraser University.

3.4. Independent Variables

There were three conditions in this study: Self-Questioning (SQ), Self-Explanation (SE), and Blank Note (BN). At the top of each of the three pages of Chapter 5, participants in all three conditions were provided the same generic instruction: *Study the page as you would normally. Each time you come across the word 'ANNOTATE' in the text, read and select the instruction next to it as you saw in the nStudy tutorial. This will highlight the text, and cause a note to open.* Following this generic instruction, each condition had unique instructions depending on the manipulation. Prompts to "Annotate" and open a note template were placed in the same positions throughout the Chapter 5 text in each condition.

3.4.1. Self-Questioning

In the self-questioning condition, participants were instructed as follows:

In the note, create a question about the content. Try to generate questions that make you think about why or how a certain technique is used, how concepts link together, or how concepts in the text link to information you already knew.

Here are question stems you will see in the text:

- What is the difference between ____ and _____?
- How does _____ affect ______?
- What would happen if ____?
- How is _________ related to ________ that I studied earlier?
- Why is _________ important?
- How _______?
- Why _________?

Since previous studies showed question prompts led to increased knowledge construction (King, 1994), these question stems were adapted from King (1994) except for the last two ("How ___? and "Why ___?") which were included because Rosenshine et al. (1996) proposed that prompts with generic question stems would be effective to students' self-generation of questions during learning. These latter two question stems allow learners to form content-based questions as well as experience-based questions that activate prior knowledge. All of these question stems could apply to any of the prompts instructing participants to self-question in the chapter.

SQ participants read Chapter 5 and, each time a prompt appeared in the text to create a question relating to the adjacent section of the content, participants selected text which opened an nStudy note template (Figure 3.4.1) containing all the question stems. Participants could select one of the question stems and fill in the blanks to generate a question. Each question stem was "controlled" by a radio button to ensure a participant could only select one stem to complete. As in previous studies (Mostow & Chen, 2009; Andre & Anderson, 1978; Duell, 1977), participants were not told to respond to the questions they had created, and no space was provided in the template to fill out a response. This avoided potential confounding of question generation with retrieval practice. Nonetheless, I could not track whether a student responded to their self-generated question, either mentally or written on a paper while studying.

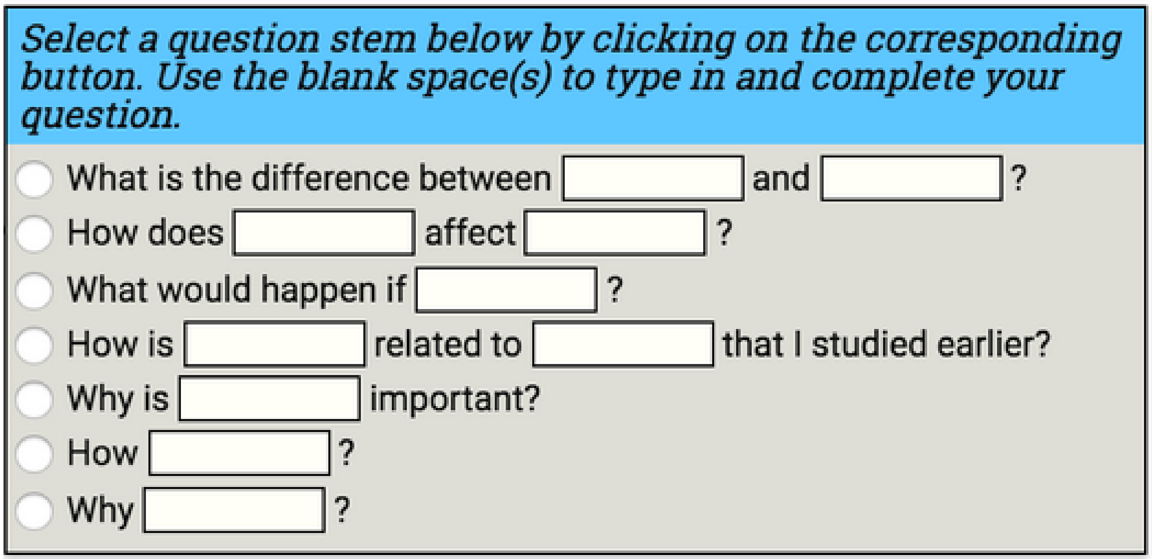


Figure 3-1. Self-questioning note template

3.4.2. Self-Explanation

In the self-explanation condition, participants were instructed as follows:

> In the note, explain the content as instructed. An explanation describes a process or a cause and its effect. **Create explanations that help you understand what you're studying.** If you need more space, the text box will expand as you type.
>
> Here is an example of a student explaining biology content to herself after reading a text:
>
> *"Substances (including vitamins, minerals, amino acids and glucose), are absorbed from the digestive system and transported to the cells. So that's why it's important to eat a balanced diet, or else your cells won't get the right vitamins, minerals amino acids, and glucose."*

SE participants read Chapter 5 and, each time a prompt appeared in the text to create a self-explanation relating to adjacent content, selected the line, which opened up an nStudy note template (Figure 3.4.2). Participants typed their explanation in the template. The example of self-explanation given to participants in the instructions was adapted from an example of self-explanation provided by Chi et al. (1994). Emphasis was placed on “you” in the instructions with underlining and bold text to ensure that students generated self-explanations, as opposed to explanations for an instructor, experimenter, or other outside reader.

Figure 3-2. Self-explanation note template

3.4.3. Blank Note

In the blank note condition, participants were instructed as follows:

> You can use the note to type out anything you might be thinking about related to the section of text you have just read.

BN participants were to read Chapter 5, and each time a prompt appeared in the text to “Create a note,” selected the line, which opened an nStudy note template (Figure 3.4.3). Participants typed their notes in the template.

Figure 3-3 Blank note template

Appendix D contains the three versions of the Chapter 5 FHS content, showing placement and wording of the prompts in each of the three conditions. Table 3.1 below compares prompts in the three conditions, placed at the same point in the text in each condition.

Table 3-1. Prompts given to students in Self-Questioning, Self-Explanation, and Blank Note conditions

Self-Questioning	Self-Explanation	Blank Note
ANNOTATE: Create a question about the Intermittent Auscultation criteria.	ANNOTATE: Explain the Intermittent Auscultation criteria.	ANNOTATE: Create a note.
ANNOTATE: Create a question about the indications for assessing Fetal Heart Rate.	ANNOTATE: Explain the indications for assessing Fetal Heart Rate.	ANNOTATE: Create a note.
ANNOTATE: Create a question about one of the preceding criteria.	ANNOTATE: Explain one of the preceding criteria.	ANNOTATE: Create a note.
ANNOTATE: Create a question about the use of IA and epidurals after reading the preceding five points.	ANNOTATE: Explain how your knowledge of the use of IA and epidurals changed after reading the preceding five points.	ANNOTATE: Create a note.
ANNOTATE: Create a question relating auscultation during the latent phase to what you know about frequency of auscultation.	ANNOTATE: Explain auscultation during the latent phase, relating it to what you know about frequency of auscultation.	ANNOTATE: Create a note.
ANNOTATE: Create a question about how non-electronic auscultation devices work in clinical practice.	ANNOTATE: Explain how non-electronic auscultation devices work and when you might use them in clinical practice.	ANNOTATE: Create a note.
ANNOTATE: Create a question about the importance of Leopold's Maneuvers.	ANNOTATE: Explain why it is important to conduct Leopold's Maneuvers.	ANNOTATE: Create a note.
ANNOTATE: Create a question about the characteristics which can't be assessed by IA.	ANNOTATE: Explain why these characteristics can't be assessed by IA.	ANNOTATE: Create a note.
ANNOTATE: Create a question about one of the situations of abnormal FHR.	ANNOTATE: Choose one of the situations of abnormal FHR above and explain it.	ANNOTATE: Create a note.
ANNOTATE: Create a question about the assessments done after detecting abnormal fetal heart rate (FHR).	ANNOTATE: Explain the assessments done after detecting abnormal fetal heart rate (FHR).	ANNOTATE: Create a note.
ANNOTATE: Create a question about the assessments done after detecting abnormal fetal heart rate (FHR).	ANNOTATE: Explain the strategies for promoting implementation of IA.	ANNOTATE: Create a note.

3.5. Design

This study has a 3 x 1 design with one independent variable, at three levels (SE, SQ, BN).

3.6. Procedure

Part 1: Recruitment and Consent

When a workshop was scheduled, instructors at hospitals, health authorities, and educational institutions email Perinatal Health Services BC listing their students' contact information and requesting access to the FHS Manual online for their students to study before the workshop. After an instructor requested access, I sent the instructor a recruitment email explaining my study. If the instructor agreed to participate, the list of students' email addresses was forwarded to me via Perinatal Health Services BC. I then emailed each student explaining the study and providing a link to an online consent form. All recruitment and consent was done online over six months from 2016 September to 2017 February.

Students completing the consent form could choose to participate in the study or not. If not, I registered them in the online FHS Manual as usual. Those who consented to participate in the study were registered into one of the duplicated versions of the online FHS Manual corresponding to the experimental condition to which they were randomly assigned: SQ, SE, or BN. Once randomly assigned, I emailed each participant with log-in information to the FHS Manual, including instructions on how to download the nStudy extension required to study the manual. Participants were sent two reminder emails to download nStudy and complete the FHS study, one month and two months after being sent initial instructions to complete the research study.

Part 2: Studying and Annotating

Participants were free to study at times of their choosing on a device and in a location of their choosing. After studying Chapters 1 to 4 on the CESEI learning management system, participants reached Chapter 5 and were automatically logged into nStudy. Here, they were redirected to a FluidSurveys online page presenting the demographic questionnaire. On the following page they took the prior knowledge test, and on the third page, a nStudy tutorial. Once participants completed these three pages, they were redirected via a link back to Chapter 5 on the CESEI system to the version of the chapter matching their experimental group. Participants were instructed: *Try to study Chapter 5 of the FHS Manual in one session.* They then studied Chapter 5 content as they wished, taking as long as they desired. The text (Appendix D) included prompts to remind participants to select text and create notes as they studied. nStudy collected time-stamped data recording whether participants opened a note at each prompt and the content of notes.

Once participants finished studying Chapter 5, a link directed them to the online posttest. First, participants answered questions about their familiarity with online technology as an interpolated task designed to enhance accuracy of the measure of longer-term recall as opposed to short-term recall. Following these two pages of questions, the posttest was presented. Each prompted recall question was shown on a separate page without an option to go back to previous pages. Once participants answered all 18 prompted recall questions, a link redirected them back to the CESEI system, and logged them out of nStudy. This marked the end of the study, and participants were free to study the rest of the FHS manual as they wished.

3.7. Data Analysis

Participants' prior knowledge was determined by counting the number of items they checked in the prior knowledge test, ranging from 0 to 9. Due to an error in formatting the test on FluidSurveys, some participants could bypass the prior knowledge test after checking off whether they knew each of the 9 items (choosing "Yes" or "No"), but before filling in the blank space to describe the concepts they had checked as known. Their prior knowledge score was the number of concepts checked “Yes.” For those participants who entered information in the blank space, additional analyses compared items they checked off as knowing to the concepts they described in the space provided. The intent was to measure agreement between what participants believe they know and the accuracy of their descriptions of the concepts they believe they know.

Posttest responses were scored using a scoring rubric (Appendix F) with a maximum possible score of 62. For further analysis, posttest scores were divided into intentional learning relating to information in the text for which participants were prompted to generate a note with a maximum 39 points, and incidental learning relating to information in the text which was not prompted for annotations with a maximum 23 points. These intentional and incidental learning scores were converted into percentages.

A co-rater and I trained to score responses to achievement posttest items on a sample of one participants’ data, then independently scored 6 randomly drawn participants' responses using the scoring rubric for posttest achievement items.

Data from FluidSurveys was combined with the nStudy data output by matching user_ids to form a complete set of data for each participant. Those who disregarded nStudy (either because they had technical difficulties and decided not to use it, or because they did not download it in the first place), were excluded from the study.

There were 12 instances distributed throughout the chapter where participants were prompted to create an annotation: a self-explanation, a self-question or a blank note. In some cases, participants did not create all 12 annotations. Thus, an annotation score was given to each participant, counting total number of annotations completed. The variable ‘annotation completed’ was operationally defined as: the participant selected a prompt in Chapter 5, which opened a note template, and entered text in the space provided. The maximum possible score was 12.

In addition, an annotation quality score was calculated. In the self-explanation (SE) condition, a score of 0 was given to each instance where there was no text entered or nonsense text was entered. A score of 1 was given if the self-explanation contained copy-and-pasted text from the source content. A score of 2 was given if participants self-explained using paraphrased content from the source, and a score of 3 was given for self-explanations that added outside knowledge, e.g. personal experience, prior knowledge, or other information not in the source.

In the self-questioning (SQ) condition, a score of 0 was given if participants left the template incomplete, blank or entered nonsense in the question blanks. A score of 1 was given for questions created that contained text copied from the source or questions created which can be answered just by reading the content. A score of 2 was assigned to questions containing added outside information or questions which could not be answered fully based on information presented in the chapter.

In the blank note (BN) condition, a score of 0 was given for those who didn’t complete the template or left it blank. A score of 1 was awarded for text copy-pasted into the space, a score of 2 was given for notes containing paraphrased content, or questions created with paraphrased content, and a score of 3 was given to notes containing outside knowledge (e.g. personal experience, prior knowledge, other information).

For each participant, a score was given for each of the 12 annotation opportunities based on quality, then totalled and divided by the maximum score possible for quality (36 for SE and BN and 24 for SQ), and multiplied by 100, to create a percentage for each participants’ annotation quality.

In the interpolated task (Appendix E), participants described their general technology skills, choosing one of “Not at all familiar,” “Slightly familiar,” Somewhat familiar,” Moderately familiar,” or “Extremely familiar.” These ratings were converted to a technology skill rating, ranging from 1 = Not at all familiar to 5 = Extremely familiar. The second question asked participants to describe their familiarity with online learning environments using the same option scale as the first question. Once again, these ratings were converted to an online learning environment familiarity rating, from 1 = Not at all familiar to 5 = Extremely familiar. These findings are reported in Appendix G.

Chapter 4.

Results

Although 160 recruits completed the consent form and indicated they would participate in the study, uptake was less than 50%. A total of 64 participants accessed the FHS Manual on the CESEI learning management system and completed the pre-test and post-test after reading Chapter 5 content (24 in SE condition, 26 in SQ, 14 in BN). Data from FluidSurveys was matched to nStudy data to confirm participants who completed the pre-and posttests also completed annotations on nStudy. Those who disregarded prompts were removed from the study. The final sample was 45: 17 in SE, 21 in SQ, and 7 in BN with complete data for analyses.

The sample consisted of 44 females and 1 male. The average age was 30 years. In a questionnaire of technology skills, participants on average rated themselves as “moderately familiar” with general technology skills and “moderately familiar” with online learning environments. Appendix G displays complete information about participants’ professions, previous experience in fetal health surveillance, reported general technology skills, and familiarity with online learning environments.

The psychometric reliability of posttest items was determined by calculating Cronbach’s coefficient of internal consistency. Across the three conditions, for the 18 posttest items, Cronbach’s alpha was $\alpha = .67$.

To investigate uptake, a chi-square test was calculated to test whether participation in the research study was equal for participants randomly assigned to each of the three conditions. The number of people who participated in the study did differ statistically detectably by condition, $\chi^2(2, N = 160) = 6.19$, $p < .05$. Those in the blank note prompt condition were less likely to participate in the research study than those participants in the SE or SQ conditions.

Table 4.1 displays descriptive statistics across all three conditions for the pretest, posttest, annotation quality score (percentage) and number of annotations completed.

Table 4-1 Mean, Standard Error, and Standard Deviation of Prior knowledge task (Pretest) scores, Posttest scores, Annotation quality scores, and Number of annotations completed across all conditions (N = 45)

	Mean	Standard Deviation	Standard Error
Pretest	4.60	3.58	0.53
Posttest	37.78	8.63	1.29
Annotation Quality (%)	61.02	19.75	2.94
Annotations Completed	10.56	2.55	.38

The items participants checked as knowing in the prior knowledge task were matched to the text descriptions they entered in the space provided. For those participants who entered text, 75.6 % of the concepts they described in the space provided matched the concepts they had checked as knowing.

Table 4.2 displays descriptive statistics in each condition for prior knowledge and posttest scores. Data were examined for normality of distributions and outliers. Pretest skewness and kurtosis for SQ and SE were ≤ ± 1.6 and ≤ ± 2.8 for BN pretest. Posttest skewness and kurtosis values were ≤ ± 1 for SQ, SE, and BN. These statistics suggest scores do not appreciably differ from the expected values of for a normal distribution.

Table 4-2 Mean, Standard Deviation, and 95% Confidence Interval for Mean of Pretest scores and Posttest scores in SQ, SE, and BN Groups (N = 45)

Condition	SQ n = 21				SE n = 17				BN n = 7			
	M	SD	95% Confidence Interval for Mean		M	SD	95% Confidence Interval for Mean		M	SD	95% Confidence Interval for Mean	
			Lower	Upper			Lower	Upper			Lower	Upper
Pretest Score	4.52	3.41	2.97	6.08	4.59	3.55	2.76	6.42	4.86	4.60	.60	9.11
Posttest Score	35.07	7.39	31.71	38.44	41.44	8.19	37.23	45.65	37.00	10.99	26.84	47.16

Figure 4.1 displays box and whisker plots for each of the three conditions for pretest scores (0–9). The ends of the box are the upper and lower quartiles; the box spans the interquartile range. The bold horizontal line in each box signifies the median, and the horizontal lines above and below the box are the maximum and minimum values, respectively. Figure 4.2 displays box and whisker plots for each of the three conditions in posttest scores (0–62).

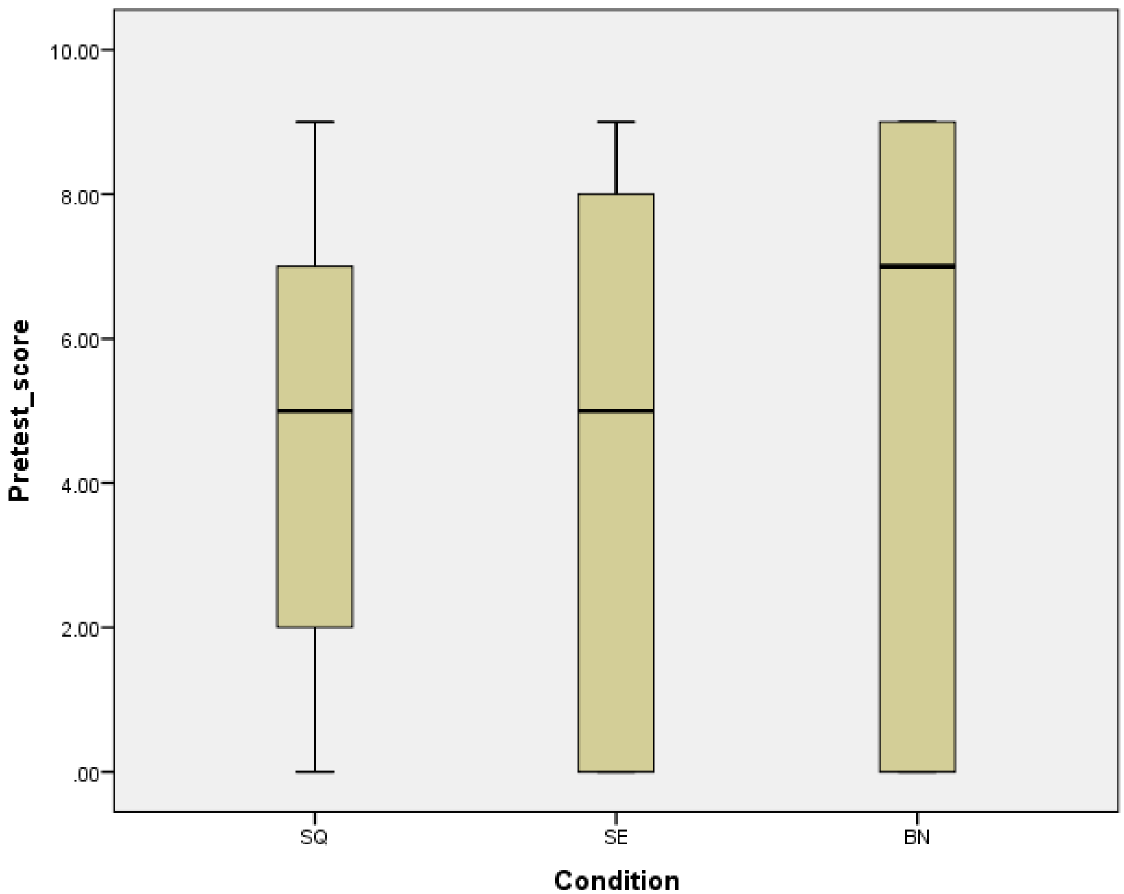


Figure 4-1 Box and whisker plot of Pretest scores in SQ, SE, and BN Groups

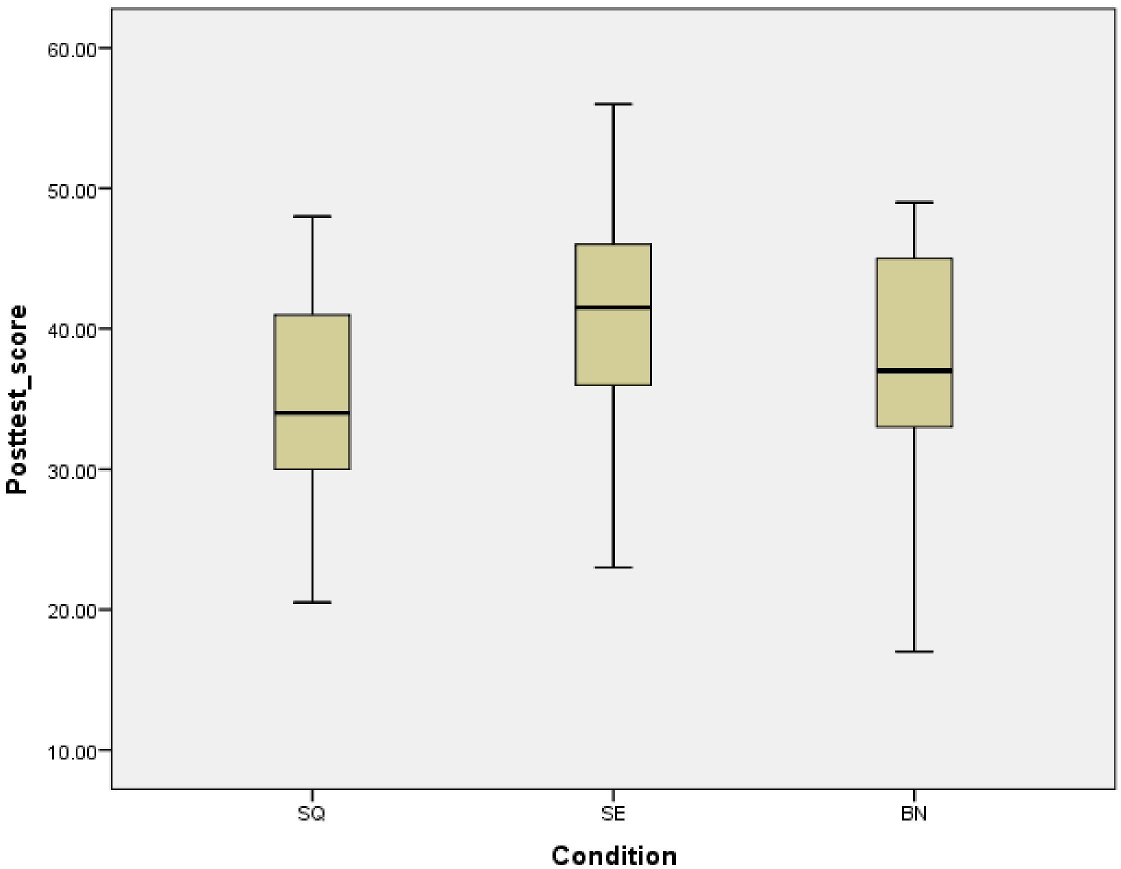


Figure 4-2 Box and whisker plot of Posttest scores in SQ, SE, and BN Groups

At pretest, there appeared to be no identifiable differences across conditions in participants' self-described knowledge of chapter content, as evident by the box and whisker plots and descriptive statistics in Table 4.1. It should be acknowledged these statistical tests have low power due to small and unequal sample sizes. From the box and whisker plots, the BN group do not seem to be identifiably differently from the SQ group in posttest score.

Through personal communication with my supervisor, we considered it prudent not to include the BN condition in statistical analyses due to the small sample size (n = 7). All further analyses use the SQ (n = 21) and SE (n = 17) groups, discarding the BN condition.

Box and whisker plots for pretest scores suggest considerable variability in the SE group compared to the SQ group. However, the two groups show similar average levels of competence on the prior knowledge task.

To ensure objective scoring of posttest achievement items, a co-rater and I independently scored 6 randomly drawn participants using the posttest scoring rubric. The correlation over scores between my scoring of posttest items and the co-rater's scores for 6 randomly drawn participants' responses was r = .943.

An independent samples t-test compared pretest scores and posttest scores between the two conditions, SE and SQ (Table 4.3). Levene's test for equality of variances was not statistically detectable, thus the assumption is met that the error variance of both the pretest and posttest scores is not different across groups (heterogeneity of variance is not a concern).

Table 4-3 Results of t-test for Posttest score by Condition

	Prompt Condition									
	SQ			SE						
	M	SD	n	M	SD	n	95% CI for Mean Difference	t	df	Sig. (2-taied)
Pretest score	4.52	3.41	21	4.59	3.55	17	-2.37, 2.24	-0.057	36	.955
Posttest score	35.07	7.39	21	41.44	8.19	17	-11.50, -1.24	-2.52*	36	.016

* $p < .05$. *Note*: CI = Confidence Interval

The t-test in Table 4.3 indicated there is no statistically detectable difference in pretest scores between SQ and SE, $t(36) = 0.057$, $p = 0.955$. Participants in both groups started out with similar levels of prior knowledge. There was a statistically detectable difference in posttest scores between SE and SQ. Those generating self-explanations (M = 41.44, SD = 8.19) performed better on the posttest than those who generated self-questions (M = 35.07, SD = 7.39), $t(36) = 2.52$, $p = 0.016$.

A t-test was calculated to examine possible differences in the number of annotations completed by each group using the annotation score (Table 4.4). There was not a detectable difference between SQ (M = 10.90, SD = 2.02) and SE (M = 10.47, SD = 3.10); $t(36)$=.520, p =

.606. Participants in the SE and SQ groups produced similar numbers of annotations. Annotation quality was also examined (Table 4.4). It appeared variability in quality was much greater in the self-explanation condition and this was confirmed by Hartley's $F_{max} = 4.80$ ($p < .01$. A t-test calculated with allowance for heterogeneous variances indicated there was not a statistically detectable difference in annotation quality between SQ and SE groups, $t(21)=.053$, $p = .958$.

Table 4-4 Results of t-test for Annotation Score and Annotation Quality by Condition

	Prompt Condition									
	SQ			SE						
	M	SD	n	M	SD	n	95% CI for Mean Difference	t	df	Sig. (2-taied)
Annotation Score	10.90	2.02	21	10.47	3.10	17	-1.26, 2.13	0.520	36	.606
Annotation Quality	61.71	12.61	21	62.09	27.64	17	-14.07, 13.30	-0.057	36	.955

p < .05. *Note*: CI = Confidence Interval

Correlations were calculated between pretest scores, posttest scores, and annotation quality in the SE and SQ groups (Table 4.5). Pretest and posttest scores were statistically detectably correlated $r(36) = .47$, $p = .003$. Annotation quality was inversely related to pretest score $r(36) = -.225$, $p = .174$, though not at a traditionally statistically detectable level.

Table 4-5 Correlations between Pretest score, Posttest score, and Annotation Quality (%) for SE and SQ groups (N = 38)

		Pretest Score	Posttest Score	Annotation Quality
Pretest Score	Pearson Correlation	1	.470**	-.225
	Sig. (2-tailed)		.003	.174
Posttest Score	Pearson Correlation	.470**	1	.145
	Sig. (2-tailed)	.003		.384

**. Correlation is significant at the 0.01 level (2-tailed).

Although all the question stems were designed to be used at any one of the twelve prompts for self-questioning throughout the chapter, some question stems were more popular than others at different prompts. Table 4.6 shows frequency of overall use of each question stem.

Table 4-6 Frequency of overall use of each question stem by SQ participants (n=21) across 12 prompts

	Question Stem	**Frequency across 12 prompts**
1	What is the difference between ____ and _____?	27
2	How does _____ affect ______?	39
3	What would happen if ____?	25
4	How is _________ related to ________ that I studied earlier?	3
5	Why is _________ important?	43
6	How _______?	39
7	Why _________?	39

Stem 4 "How is _________ related to ________ that I studied earlier?" was rarely used by participants. Stem 5 "Why is _________ important?" was used most often, followed by a three-way tie by "How __?" "Why __?" and "How does _____ affect ______?"

Table 4-7 Multivariate test comparing SQ and SE incidental and intentional scores

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.963	453.637[b]	2.000	35.000	.000
	Wilks' Lambda	.037	453.637[b]	2.000	35.000	.000
	Hotelling's Trace	25.922	453.637[b]	2.000	35.000	.000
	Roy's Largest Root	25.922	453.637[b]	2.000	35.000	.000
Condition	Pillai's Trace	.245	5.680[b]	2.000	35.000	.007
	Wilks' Lambda	.755	5.680[b]	2.000	35.000	.007
	Hotelling's Trace	.325	5.680[b]	2.000	35.000	.007
	Roy's Largest Root	.325	5.680[b]	2.000	35.000	.007

a. Design: Intercept + Condition
b. Exact statistic

A two-group between-subjects multivariate analysis of variance was calculated to compare SQ and SE participants' performance on incidental (umprompted content) and intentional (prompted content) items of the posttest (Table 4.7). Box's test of equality of the variance-covariance matrices $p = .348$ was not statistically detectable, indicating the observed covariance matrices of the dependent variables were not different across groups, Wilks' Lambda evaluates all multivariate effects simultaneously. The composite dependent variate was statistically detectably affected by condition, $\Lambda = .755$, $F(2, 35) = 5.68$, $p = .007$. Univariate ANOVAs were calculated for each dependent measure separately to investigate the locus of the statistically detectable multivariate effect (Table 4.8). A statistically detectable univariate effect was associated with intentional scores, $F(1, 36) = 9.64$, $p = .004$, $\eta^2 = .211$. Participants in the SE condition had much higher scores on intentional (prompted content) posttest items ($M = 65.99$, $SD = 14.75$) than participants in the SQ condition ($M = 51.28$, $SD = 14.33$) but there was no statistically detectable difference on incidental items ($p = .756$).

Table 4-8 Tests of Between-Subject Effects for Intentional and Incidental Scores

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Incidental_Percent	21.599[a]	1	21.599	.098	.756
	Intentional_Percent	2031.734[b]	1	2031.734	9.640	.004
Intercept	Incidental_Percent	171362.759	1	171362.759	776.821	.000
	Intentional_Percent	129198.764	1	129198.764	613.002	.000
Condition	Incidental_Percent	21.599	1	21.599	.098	.756
	Intentional_Percent	2031.734	1	2031.734	9.640	.004
Error	Incidental_Percent	7941.415	36	220.595		
	Intentional_Percent	7587.500	36	210.764		
Total	Incidental_Percent	180836.484	38			
	Intentional_Percent	136839.250	38			

Table 4.9 displays descriptive statistics for these measures. Values for skewness and kurtosis were ≤ ± 1 for all scores. Univariate analyses reveal that on incidental items, there was no statistically detectable difference between the SE and SQ groups. For posttest intentional items there was a statistically detectable difference in scores between the SE and SQ groups. The effect size was calculated by using a pooled standard deviation of the mean ($SD_{pooled} = 14.54$), and was found to be $d = 1.01$, suggesting a moderate practical significance.

Table 4-9 Mean and Standard Deviation for Incidental Score and Intentional Score in Self-questioning and Self-explanation groups

	Prompt Condition					
	SQ			SE		
	M	SD	n	M	SD	n
Incidental Score	66.77	12.75	21	68.29	17.12	17
Intentional Score	51.28	14.33	21	65.99	14.75	17

Comments from participants were collected and provided in Appendix H. Three main themes emerged. Some participants thought the software was beneficial for annotating, others had some issues with the instructions to annotate or found it inconvenient, and some believed it hindered their studying processes. 17 participants in the SQ condition commented about the nStudy annotation software. Of these, 3 commented the software was user-friendly and helpful. 10 participants who generated self-questions found it frustrating. They found that stopping to generate questions was tedious, distracting, or impaired their mode of studying. Learners commented trying to fit content into blanks of one of the provided question stem was frustrating and detracted from learning. One learner would have preferred to create self-questions from scratch rather than choosing from a provided blank stem. Of 15 participants in the self-explanation condition who commented on the annotations, 8 were generally positive, indicating that it took a bit of time to figure out but was generally a good learning aid. One participant mentioned it forced her to stop and re-read to create better self-explanations. Another two said generating self-explanations helped them think more deeply about course material while studying.

Chapter 5.

Discussion

Overall achievement across all three conditions was *M* = 37.78 (*SD* = 8.63), out of a possible score of 62; participants received an average score of 61% on the posttest. This result is important for health professionals and their trainers. If this study module is to prepare students to participate effectively in a professional development workshop, it appears not to succeed well enough. Equipping learners to better use learning strategies may boost achievement.

From a learning science perspective, it is important to investigate uses and potential effects of the learning strategy prompts. Several factors merit exploration relating to why participants who created self-explanations while studying had higher performance on the prompted recall posttest than participants who generated self-questions. Because pretest scores were similar across conditions, prior knowledge was discounted as a possibility.

Although learners who self-explained had higher posttest performance, this is not because the SE group simply made more annotations during the study session. Both SE and SQ learners produced similar numbers of annotations. If making annotations in either the form of a self-explanation or question generation could be interpreted as effort or attentiveness, both groups seem to be applying similar effort or attentiveness to the studying task.

The uptake by students of this online study and annotation scores indicate some students in each condition were diligent in working within the condition to which they were assigned. Others decided to forego instructions to use annotations and studied as they wished. Thus, the annotation quality score is valuable in that perhaps some students who were instructed to take notes by creating questions or explanations, may not have been engaged in annotating. In other words, this score may provide more information about effort applied. Comparing annotation quality between SE and SQ, there was no statistically detectable difference, and there was a lot of variability in annotation quality within the self-explanation group. Considering participant comments (Appendix H), while several SE participants found the annotation prompts beneficial, others did not feel confident, or said they found it inconvenient and thus did not complete annotations throughout the chapter. Although not statistically detectable, annotation quality was inversely related to pretest score $r(36) = -.225$, $p = .174$. This might suggest participants with higher prior knowledge would make lower quality annotations. One possible account for this negative relation between prior knowledge and quality of notes is that participants did not feel it was necessary for them to create such elaborate annotations on information they thought they knew about. Fine-grained analysis examining the match between specific topics participants indicated they knew about, and the content quality of their annotations on that specific topic, also could help to investigate this.

It is important to note participants were instructed to generate self-explanations or self-questions at the start of the study session, and subsequently saw prompts to annotate. Prompts are not demands – learners ultimately decide whether they would like to generate an annotation. In a lab experiment situation, following such instructions may have a different uptake, and participants might feel their circumstances necessitate more deliberate note-taking, especially if the experimenter is in the same room. However, in a real-world online study such as this, where participants are busy professionals (registered nurses, medical doctors, etc.) or very busy students, it may be hard to comply fully with annotation instructions, especially if participants want just to read the chapter and move on to other obligations. Participants may experience tension about applying effort to this study considering their other duties.

As evident by comparing incidental and intentional achievement on the posttest, generating self-questions suppressed learning compared to creating self-explanations. Cohen's effect size value ($d = 1.01$) suggests a moderate to high practical significance.

Participants generating self-questions achieved less on the posttest than those generating self-explanations despite a similar level of quality in their annotations. It seems both groups are elaborating and working with information at a similar level of engagement. For self-questioning, this engagement with the text was less successful than self explaining. Setting aside the possibility this is a type I error, there must be some factors at play beyond engaging with the information that differs between self-explanation and self-questioning. Perhaps trying to explain reasoning behind a cause or the importance of certain techniques, as is prompted in the study, is more productive for promoting learning than asking questions.

However, it is important to note the data collected in this study do not support strong inferences about participants' metacognitive activity while generating self-explanations or self-questions. For example, in self-questioning, it can not be determined whether students evaluated each of the provided question stems individually, and then chose the one for which they could best create a question; or whether students thought of a question, then chose a stem which best fit it, or whether at each prompt they chose a question stem at random and filled in its blanks without considering the other stems. Similarly, for self-explanation participants' metacognition is not clearly identifiable. For example, it is not known whether they metacognitively monitored relevant prior knowledge before generating an explanation, or chose to explain the first idea which came to their mind related to the prompted content.

In the case of self-questioning, 7 possible stems were provided, of which learners were to choose one and fill in blank spaces. Although stems were chosen based on past research that found knowledge was enhanced knowledge when such stems were provided (King, 1994) or more generic stems were offered, (Rosenshine et al., 1996), stems did not work well in this study. In fact, self-questioning prompts suppressed learning compared to both non-prompted self-questioning material and self-explaining. In King's (1994) study, students were fourth and fifth graders asking questions about a science lesson. Perhaps for that age group, it is beneficial to provide more structured question stems, something that older, mature students would not prefer. College students were also provided these stems in generating self-questions (King, 1992). However, participants in King's study had low scores on the Scholastic Aptitude Test and were enrolled in a remedial reading course (King, 1992). For those learners, a question stem structure can support using the strategy whereas relatively advanced learners like those participating in the present study might find a structured template too restrictive, and would want to generate questions in a more open-ended manner.

In King's (1994) study, students received explanation training along with question generation guidance, and were to ask each other questions and respond to these questions using explanations. This might imply it was forming explanations that led to greater comprehension in King's study. However, a comparison group also received training in formulating explanations but unguided in questioning. It did not perform as well on comprehension and recall tasks. One other difference is that students were told in King (1994) that generating questions for themselves and others would be beneficial to learning. In my study, students are solely generating questions for themselves, and not using them to generate discussion with a peer. However, in King (1992), college students formed self-questions for themselves, rather than to ask peers, as in my study, and were found to have improved long-term recall compared to a summarization group and note-taking group. That study took place in a paper-and-pencil format, rather than creating questions online.

Participants in the SQ condition rarely chose to create a question relating information to things they already knew or had studied earlier. Perhaps more integration with prior knowledge when generating questions could enhance recall, as in other research (King, 1992). The most popular question stem was asking about why a certain concept was important. The two most generic question stems, (How __? and Why __?) along with the question stem connecting concepts from the text (How does _____ affect ______?) were also highly frequent. It may be that with the content presented in this manual, certain types of question stems might allow for more beneficial learning opportunities. Perhaps if more learners had chosen the question stem designed to connect information to prior knowledge (which was used considerably less frequently than all the other stems), learning would be enhanced.

Past literature has suggested self-questioning is an opportunity for students to connect new information to what they know; thus, there is opportunity to assemble information. As can be seen by the annotation quality scores, there are several participants who do that assembly at minimal levels (i.e. copy-pasting text from the content rather than incorporating outside knowledge). Even for those who did assemble information through self-questioning, they didn’t often choose the question stem that brings prior knowledge into the questions. It appears from the findings of my study that fitting information into a self-explanation schema is more beneficial than trying to create self-questions.

At the same time, self-questioning in this study might be seen as a selection task, choosing the best stem to use, rather than a purely self-generative activity, in which a participant might use a non-structured open-ended template to create a question. This contrasts with the self-explanation condition, in which participants were able to generate an explanation for themselves without any pre-determined stems.

As seen in Table 3.1, prompts in each condition were designed to be as similar as possible with respect to information participants would process, so the only difference would be the form of engagement, a question or a self-explanation.

It is important to put control of learning in the hands of the student, and to have them self-regulate and decide what learning strategies to use and when to use them (Murad et al., 2010). Thus, perhaps the act of stopping at specified points in the text to create a question disrupted learners' studying activities rather than enhancing them. Providing prompts at pre-determined points throughout the text was done to signal content suitable for the forms of annotation examined here, and that would be tested on later. However, a more free-flowing, open-ended study method might have worked better for participants in the self-questioning condition. In King (1992), college students were given similar questions stems as in my study, which was thought to aid in enhancing their learning. Again, perhaps a more experienced group of learners would not need so much guidance in their use of self-questioning strategy. However, in King (1992), students were free to form questions during studying whenever they wanted – there were no predetermined points during the lesson in which to self-question. If the learners in my study had the opportunity to choose when to create a question, or could generate self-questions without having to choose from a list of question stems, they might have been better able to use the self-questioning strategy at a level suited to their knowledge, and address gaps in that knowledge.

Like my study, King (1992, 1994) trained students in questioning strategies, and provided them with guided prompts. However, the training provided in my study consisted of a brief passage of text instructions before learners started studying the chapter, rather than a human trainer. Perhaps a more comprehensive training session, complete with modeling, scaffolding, and opportunity to practice generating questions would enhance results of self-questioning.

Descriptions participants provided in the prior knowledge task were not analyzed; their pretest score was based on the number of concepts checked rather than knowledge objectively determined. A participant's perception of what he or she knows is what drives their studying more than an independent observer's score of knowledge. This was supported by the analysis comparing items participants checked as known to text descriptions they entered in a space provided. When students checked an item they judged they knew, they could describe that item well on average 75.6% of the time.

A strength of this study is learners are potentially more motivated to master content of the online manual they studied and on which they were tested. Participants did not volunteer to learn extraneous information. Rather, they were required to study this manual as professionals who elected to enhance the health and safety of a vulnerable population. This motivation is rare in lab studies in which students study a topic chosen by researchers.

5.1. Opportunities for Further Research

Important motivational features may need to be identified and subsequently studied in this kind of research. It would be beneficial to learn what affects uptake of a study tactic in general and by those in varying professions. People were more responsive to the two new methods of note-taking strategies than the standard, familiar way of making notes. Is there something about the intervention that encourages those instructed to self-explain or self-question to engage with greater effort in the task and complete the research study compared to participants told to simply take notes with no additional learning strategy? Further research could explore learners' motivations to use different study tactics in real online learning settings.

Clearly, it is hard to have much control in a study like mine, as all instructions and material studied are accessed online at a time and in a place the participant chooses. Because participants' confidentiality was maintained and no identifiable information was linked to studying activities, it was not possible to identify which individuals completed the study and which individuals disregarded the research activities. Thus, I was unable to send specific reminder emails to those who had not completed the study, aside from the two general reminders sent to all participants who consented. Nonetheless, it is important to note that perhaps powerful findings which come about in lab studies may weaken or even disappear in the real world due to a variety of individual factors and motivational differences. Exerting more control in a lab setting to ensure a stronger scientific test may not predict the likelihood such results will occur in a real learning situation, despite the hope of many learning scientists.

Due to small sample size within each of the SE and SQ groups, analyses of covariance were not done to test whether a learner's profession (i.e., Registered Nurse, Medical Doctor, or Midwife), or educational level affected performance on the posttest. Further research might explore the differences between these distinct professions to determine whether a kind or level of training leads to more effective use of study strategies.

Future versions of the study might vary slightly by (a) allowing those in the self-questioning condition to generate a response to their question, (b) giving learners more autonomy to choose for themselves when they want to generate a question, or (c) providing a blank note template for self-questioning rather than a list of pre-determined question stems. A think-aloud protocol collect data to guide inferences about participants' metacognitive activity engaged in learning under a self-questioning regime.

5.2. Conclusion

Based on the results of this study, providing structured question stems with blank spaces at pre-determined locations in a text is not beneficial to health professionals and students studying fetal health online. Rather, it suppressed learning when measuring prompted recall compared to non-prompted material. Perhaps placement of the prompts disrupted learners' studying process as they aimed to understand information presented. Or, it could be the specific stems learners chose to use most frequently were not conducive to learning this subject matter.

For those who generated self-explanations, recall performance on the posttest was about the same for prompted material and unprompted material. Inviting learners to explain information to themselves may benefit learning such content by promoting reasoning or relating new information to prior knowledge. Research with a larger sample size might further explore the specific mechanisms health professionals use while studying and generating self-explanations.

In summary, it is interesting to note learners' perceptions of the annotation prompts and their uptake considering the sample characteristics and circumstances of nation-wide online learning. If prompts are used in this context, self-explanation rather than self-questioning prompting is recommended.

References

Andre, M. E., & Anderson, T. H. (1978). The development and evaluation of a self-questioning study technique. *Reading Research Quarterly*, 605-623.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist, 40*(4), 199-209.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523.

Berry, D. C. (1983). Metacognitive experience and transfer of logical reasoning. *The Quarterly Journal of Experimental Psychology, 35*(1), 39-49.

Bisra, K., Liu, Q., Salimi, F., Nesbit, J. C., & Winne, P. H. (2016). Inducing self-explanation: A meta-analysis. Manuscript submitted for publication.

Burnett, E. H., & Berg, P. C. (1988). Reading instruction in the schools: Improving students' critical thinking skills. The Clearing House: *A Journal of Educational Strategies, Issues and Ideas, 61*(5), 208-210.

Cano García, F., García, Á., García-Berbén, A. B., Pichardo Martínez, M. D. C., & Justicia Justicia, F. (2014). The effects of question-generation training on metacognitive knowledge, self regulation and learning approaches in Science. *Psicothema, 26*(3), 385-390.

Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science, 18*(3), 439-477.

Chumley-Jones, H. S., Dobbie, A., & Alford, C. L. (2002). Web-based learning: Sound educational method or hype? A review of the evaluation literature. *Academic Medicine, 77*(10), S86-S93.

Duell, O. K. (1977, April). *Overt and covert use of objectives of different levels.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Fukaya, T. (2013). Explanation generation, not explanation expectancy, improves metacomprehension accuracy. *Metacognition and Learning, 8*(1), 1-18.

Graffam, B. (2007). Active learning in medical education: strategies for beginning implementation. *Medical Teacher, 29*(1), 38-42.

Hausmann, R. G., & Chi, M. H. (2002). Can a computer interface support self-explaining. *Cognitive Technology, 7*(1), 4-14.

Huang, Z. (1992). *A meta-analysis of student self-questioning strategies.* Dissertation, Hofstra University (UMI Dissertation Services No. 9211456).

Janssen, T. (2002). Instruction in self-questioning as a literary reading strategy: An exploration of empirical research. *L1-Educational Studies in Language and Literature, 2*(2), 95-120.

Johnson, C. E., Hurtubise, L. C., Castrop, J., French, G., Groner, J., Ladinsky, M. McLaughlin, D., Plachta, L., & Mahan, J. D. (2004). Learning management systems: technology to measure the medical knowledge competency of the ACGME. *Medical Education, 38*(6), 599-608.

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

King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures. *American Educational Research Journal, 29*(2), 303-323.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*(2), 338-368.

Knowles, M. (1975). *Self—Directed Learning: A Guide for Learners and Teachers*. Nueva York: Cambridge. The Adult Education Company.

Larsen, D. P., Butler, A. C., & Roediger III, H. L. (2013). Comparative effects of test-enhanced learning and self-explanation on long- term retention. *Medical Education, 47*(7), 674-682.

Mayer, R. E. (2010). Applying the science of learning to medical education. *Medical Education, 44(*6), 543-549.

Mostow, J., & Chen, W. (2009, July). Generating Instruction Automatically for the Reading Strategy of Self-Questioning. In AIED (pp. 465-472).

Murad, M. H., Coto-Yglesias, F., Varkey, P., Prokop, L. J., & Murad, A. L. (2010). The effectiveness of self-directed l earning in health professions education: a systematic review. *Medical Education, 44*(11), 1057-1068.

Papadopoulos, P. M., Demetriadis, S. N., Stamelos, I. G., & Tsoukalas, I. A. (2011). The value of writing-to-learn when using question prompts to support web-based learning in ill-structured domains. *Educational Technology Research and Development, 59*(1), 71-90.

Perinatal Services BC (2014, November). Online manual developed for Fetal Health Surveillance. SOGC news. Retrieved from http://old.sogc.org/newsletter/november-2014/online-manual-developed-fetal-health-surveillance/

Roediger III, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*(3), 181-210.

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

Ruiz, J. G., Mintzer, M. J., & Leipzig, R. M. (2006). The impact of e-learning in medical education. *Academic Medicine, 81*(3), 207-212.

Sandars, J., & Schroter, S. (2007). Web 2.0 technologies for undergraduate and postgraduate medical education: an online survey. *Postgraduate Medical Journal, 83*(986), 759-762.

Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers & Education, 46*(4), 426-445.

Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Granott, & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning*, 31-58. Cambridge: Cambridge University Press.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes*, *21*(3), 255-287.

Trevors, G., Duffy, M., & Azevedo, R. (2014). Note-taking within MetaTutor: interactions between an intelligent tutoring system and prior knowledge on note-taking and learning. *Educational Technology Research and Development*, *62*(5), 507-528.

Westera, J. & Moore, D.W. (1995). Reciprocal teaching of reading comprehension in a New Zealand high school. *Psychology in the Schools,* 32, 225–232.

Wong, R. M., Lawson, M. J., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learning and Instruction, 12*(2), 233-262.

Wong, B. Y. (1985). Self-questioning instructional research: A review. *Review of Educational Research*, *55*(2), 227-268.

Zehry, K., Halder, N., & Theodosiou, L. (2011). E-Learning in medical education in the United Kingdom. *Procedia-Social and Behavioral Sciences,* 15, 3163-3167.

Appendix A.

Distribution of Participant Professions

Table A1. Frequency of Participants' Profession in each Condition

	Self-Questioning (n = 21)	Self-Explanation (n = 17)	Blank Note (n = 7)	Total Across Conditions
Registered Nurse	9	8	2	19
Medical Doctor	2	0	1	3
Midwife	1	2	2	5
Nursing Student	4	0	1	5
Medical Student	1	2	1	4
Student Midwife	4	5	0	9

Appendix B.

nStudy Tutorial

Below is the nStudy tutorial given to participants in the SE group. The tutorial differs for SQ and BN groups only in that they see different ‘ANNOTATE’ instructions (rather than “Explain…”) and the version of the nStudy template corresponding to their condition.

Sample nStudy tutorial:

nStudy is an online learning tool. It helps you organize information you study and manage that information. It also keeps track of your work within specific pages.

nStudy requires you to install an extension to your Chrome web browser, ensure that you have followed the steps to do so (you should have been emailed a document: How to download nStudy extension). Please email dsamadi@sfu.ca if you are having trouble downloading nStudy.

Try to study Chapter 5 of the FHS Manual in one session.

Each time you come across the word ‘ANNOTATE’ in the text, you will be asked to create a nStudy note associated with that text.

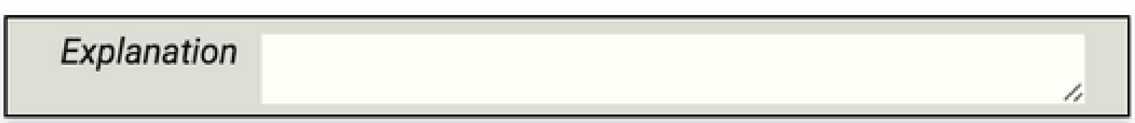
To create a Note in nStudy:

1. Select only the 'ANNOTATE' line of the text with your cursor:

ANNOTATE: Explain the following…

ANNOTATE: Explain the following…

2. The note will open up to the right of your screen and you can start typing in it. If the note does not open, try selecting the text again.



IMPORTANT: In this study, you are not able to review notes you have created. The focus is on the effect of creating notes. Thus, once you click outside the note window, you will not be able to access or view your note again.

Therefore, please fill out the note template, as instructed in each "ANNOTATE" prompt, and then once you have completed this task, you can click outside the note window to close the note. If you accidentally click outside of the note window before you are finished filling out your note, you can just re-select the text and create a new note.

*** If you have already selected the 'ANNOTATE' text, it turned blue, but a note template did not pop up:

You can double-click on the blue-highlighted ANNOTATE text, and the note should open. Alternatively, you can re-select already highlighted text and a new note will open.

Appendix C.

Demographic Questionnaire and Prior Knowledge test (Pretest)

Demographic Questionnaire:

Age: ______

Gender:

○ Male
○ Female
○ Other: ____

What is your highest educational degree (e.g. High school, BSc/BA, MA, PhD) and in which discipline? __________________________

Please indicate your profession:

○ RN
○ Nursing student

○ MD

○ Medical Student

○ Midwife

○ Student Midwife

○ Other: ________

Have you taken a fetal health surveillance course in the past? If so, please list which course(s) you have completed below:

____

Do you have prior experience working with pregnant woman, doing fetal health surveillance? If so, please indicate approx. number of months.
○ Yes
○ No

Number of months doing fetal health surveillance:

____

Pretest

Here are some concepts related to Chapter 5 of the FHS Module.

Choose "Yes" for the ones that you think you could describe well to a classmate.

In the space below the list, give a brief description of the concepts you think you know about.

Key Concepts	**Yes**	**No**
When is intermittent auscultation used	☐	☐
Risks to consider when using intermittent auscultation	☐	☐
Benefits of intermittent auscultation	☐	☐
Frequency of auscultation during different phases of labour	☐	☐
Examples of auscultation devices	☐	☐
Interpretation of normal fetal heart rate	☐	☐
Interpretation of abnormal fetal heart rate	☐	☐
When to perform Leopold’s Maneuvers	☐	☐
How to perform Leopold’s Maneuvers	☐	☐

In the space below, give a brief description of the concepts you think you know about.

Appendix D.

Studying Text

Below is the Chapter 5 FHS Manual text with prompts. To avoid redundancy in including the text three times for each condition, the text is displayed once, showing each of the three prompt types at the location they appeared in the text for learners. Participants saw the prompt of whichever condition they were assigned. ANNOTATE prompts have been highlighted here to ease viewing, however they were not highlighted in the original text which participants read.

INTERMITTENT AUSCULTATION CRITERIA

● The presence of practitioners experienced in the technique of auscultation, the palpation of contractions, and the auditory recognition of pertinent FHR changes.

Self-explanation (SE): ANNOTATE: Explain the Intermittent Auscultation criteria.
Self-questioning (SQ): ANNOTATE: Create a question about the Intermittent Auscultation criteria.
Blank Note (BN): ANNOTATE: Create a note.

● The existence of a protocol addressing the technique, frequency of assessment and response (as per SOGC recommendations, 2007).

INDICATIONS FOR AUSCULTATION

1. Healthy Term Women without risk factors for adverse perinatal outcomes: at initial assessment in triage and throughout labour
2. Other Common Indications for Auscultation
Assess FHR before:
● initiation of labour-enhancing procedures (e.g. amniotomy)
● administration of medications
● administration or initiation of analgesia/anaesthesia
● transfer or discharge of patient

SE: ANNOTATE: Explain the indications for assessing Fetal Heart Rate.
SQ: ANNOTATE: Create a question about the indications for assessing Fetal Heart Rate.
BN: ANNOTATE: Create a note.

Assess FHR after:

● admission of patient
● artificial or spontaneous rupture of membranes
● vaginal examinations
● abnormal uterine activity patterns (e.g. increased basal tone or tachysystole)
● any untoward event during labour (e.g. maternal hypotension, bleeding)
● administration or initiation of analgesia/anaesthesia

SE: ANNOTATE: Explain one of the preceding criteria.
SQ: ANNOTATE: Create a question about one of the preceding criteria.
BN: ANNOTATE: Create a note.

3. Intermittent Auscultation in Specific Clinical Situations

Epidural Analgesia

● "In general, epidural or spinal anaesthesia in the absence of maternal hypotension or uterine hypertonus causes minimal changes in the FHR" (Capogna, 2001).
● The use of IA is appropriate after initiation of regional analgesia in healthy term women without risk factors.
● It is reasonable to increase the frequency of auscultation to every 5 minutes for 30 minutes after the initial epidural dose and after any additional bolus top-ups of concentrated local anaesthetic (SOGC, 2007)
● Patient-controlled epidural analgesia (PCEA) uses a more diluted local anaesthetic and opioid solution than for the traditional epidural infusion with intermittent top-ups. For this reason, PCEA allows ambulation in labour and maternal hypotension is rarely a problem. Therefore, IA is acceptable and should be performed according to usual protocols (and not after each self-administered top-up).
● If maternal hypotension occurs and is a persistent problem, continuous electronic fetal monitoring should be initiated.

SE: ANNOTATE: Explain how your knowledge of the use of IA and epidurals changed after reading the preceding five points.
SQ: ANNOTATE: Create a question about the use of IA and epidurals after reading the preceding five points.
BN: ANNOTATE: Create a note.

Post-Dates / Post-Term

● Up to 41+3 weeks gestation, spontaneous labour with no risk factors: IA is preferred
● After 41+3 weeks gestation, with normal non-stress tests (NSTs) and amniotic fluid volume:

IA is preferred
● After 42 weeks gestation (post-term pregnancy): EFM is preferred due to increased risk of adverse fetal outcome

Preterm Labour

● Due to increased incidence of adverse outcome and other pathologies, EFM is recommended in preterm fetuses below 36 weeks gestation

Trial of Labour / Vaginal Birth After Cesarean (VBAC)

● Continuous EFM is recommended

RECOMMENDED FREQUENCY OF AUSCULTATION (SOGC, 2007)

1. During the Latent Phase

● There is not a lot of data on which to base a recommendation for fetal heart observation during the latent phase of labour. Optimally, most women will be at home with family support during this period.

● If the woman is in hospital during this phase:

- ■ Low-risk: at time of assessment and approximately every hour

- ■ When a change in her condition occurs (such as rupture of the membranes, development of bleeding or other concerning clinical events), fetal surveillance using the most appropriate method should be instituted, as per SOGC guidelines (2007).

SE: ANNOTATE: Explain auscultation during the latent phase, relating it to what you know about frequency of auscultation.
SQ: ANNOTATE: Create a question relating auscultation during the latent phase to what you know about frequency of auscultation.
BN: ANNOTATE: Create a note.

2. During Active Labour and Second Stage

● Every 15 – 30 minutes in the active phase of the first stage of labour, and before the onset of pushing in the second stage of labour
● Every 5 minutes in the **active second stage** of labour, once the woman has begun pushing.

AUSCULTATION DEVICES (Adapted from Fischbeck-Feinstein, Sprague & Trépanier, 2008)

Non-electronic auscultation devices (e.g. **fetoscope** or **Pinard-type stethoscope**) allow practitioners to hear the actual fetal heart sounds associated with the opening and closing of ventricular valves via bone conduction (Goodwin, 2000). A **Doppler** device uses ultrasound technology (similar to the external ultrasound transducer of the electronic fetal monitor) to detect fetal heart movements such as the moving heart walls or valves. This ultrasound device then converts that information into a sound that represents the cardiac movements. The two methods obtain information differently but are both appropriate in most auscultation clinical situations.

SE: ANNOTATE: Explain how non-electronic auscultation devices work and when you might use them in clinical practice.
SQ: ANNOTATE: Create a question about how non-electronic auscultation devices work in clinical practice.
BN: ANNOTATE: Create a note.

A fetoscope or Pinard-type stethoscope can be used to detect the FHR baseline, rhythm, and changes from the baseline, and can identify the presence of an irregular rhythm. Doppler devices can also detect the FHR baseline, rhythm, and changes from the baseline. However, when a dysrhythmia is identified, further assessment with other methods (e.g. ultrasound, echocardiography) may be necessary to determine the type of dysrhythmia or to rule out artifact (SOGC 2007).

Popular auscultation devices include dopplers that display the FHR digitally, potentially eliminating the need to count the FHR, and water-immersible dopplers that are safe for use with women labouring in the bathtub or shower. Sometimes, practitioners use the external ultrasound transducer of the fetal EFM as an auscultation device *without generating a tracing* (i.e. the paper recorder is turned off), rather than using a hand-held Doppler device. Some institutions' electronic fetal monitors are linked to computerized systems which automatically archive the data obtained, whether the tracing is being printed or not. When performing intermittent auscultation using the external transducer of these monitors, the short segments of data could theoretically be retrieved for subsequent review; however the information generated would be too brief and incomplete to be interpretable as "EFM tracings". Therefore, for IA, it is best to use a hand-held Doppler or fetoscope. Issues related to devices used should be addressed by institutional policies and procedures.

SE: ANNOTATE: Explain this last paragraph to understand why it is best to use a hand-held Doppler or fetoscope for Intermittent Auscultation.
SQ: ANNOTATE: Create a question about this last paragraph, focusing on the use of a hand-held Doppler or fetoscope for Intermittent Auscultation.
BN: ANNOTATE: Create a note.

RECOMMENDED PROCEDURE FOR FHR AUSCULTATION (Adapted from Fischbeck-Feinstein, Sprague & Trépanier, 2007)

● Palpate the maternal abdomen to identify fetal presentation and position (Leopold's Maneuvers, see next page).

● Place the Doppler over the area of maximum intensity of fetal heart sounds (usually over the fetal back or shoulder).

● Listen to hear the FHR and place a finger on mother's radial pulse to differentiate maternal from fetal heart rate.

● Establish a baseline heart rate by listening and counting between uterine contractions for a full minute (60 seconds).

● Once the FHR baseline is established auscultate the FHR immediately after a contraction for ongoing readings.

NOTE: Although both 30- and 60-second counting periods are used in practice, little evidence exists regarding best counting duration. However, a 60-second count improves accuracy as it is longer and most likely reflects the number of fetal heart beats per minute. Also, some clinicians listen through a contraction but this practice is not supported by research evidence.

● To clarify FHR accelerations and decelerations, consecutive counts for 6-second intervals (multiplied by 10) may be helpful (although this practice is not supported by research evidence).

LEOPOLD'S MANEUVERS

FIRST MANEUVER

Face the client. Palpate the upper abdomen with both hands. Note the shape, consistency, and mobility of the palpated part. The fetal head is firm, hard, and round and moves independently of the trunk. The breech (buttocks) feels softer and moves with the trunk.

SECOND MANEUVER

Moving the hands down toward the pelvis, palpate the abdomen with gentle but deep pressure. The fetal back, on one side of the abdomen, feels smooth. This is the area that provides the most audible fetal heart sounds. On the opposite side, the fetal arms, legs, and feet feel knobbly and bumpy.

THIRD MANEUVER

Place one hand just above the symphysis. Note whether the part palpated feels like the fetal head or the breech and whether the presenting part is engaged.

FOURTH MANEUVER

Face the client's feet. Place both hands on the lower abdomen, and move the fingers of both hands gently down the sides of the uterus toward the pubis. Note the cephalic prominence or brow on the opposite side of the fetal back.

SE: ANNOTATE: Explain why it is important to conduct Leopold's Maneuvers.
SQ: ANNOTATE: Create a question about the importance of Leopold's Maneuvers.
BN: ANNOTATE: Create a note.

INTERMITTENT AUSCULTATION ASSESSMENT

1. FHR characteristics that can be assessed by IA

- **Baseline fetal heart rate** (the heart rate counted for one minute between contractions, in bpm).
- **Rhythm** (regular or irregular)
- **Presence of accelerations**
- **Presence of decelerations**

2. What <u>cannot</u> be assessed by IA?

- **Baseline variability.** This is a characteristic *visually* assessed on EFM tracing. Although some practitioners believe counting the FHR for shorter and more frequent intervals may provide information about variability, there are no research data that support this practice.
- **Classification/type of deceleration.** There is no research to indicate that the deceleration patterns classified *visually* with EFM can be classified with similar terminology when they are auscultated.
- **Other *visually* identifiable FHR changes** such as sinusoidal pattern.

SE: ANNOTATE: Explain why these characteristics can't be assessed by IA.
SQ: ANNOTATE: Create a question about the characteristics which can't be assessed by IA.
BN: ANNOTATE: Create a note.

INTERPRETATION

1. Normal FHR

- Normal baseline rate range: 110 to 160 bpm
- Presence of accelerations (increases in FHR)

2. Abnormal FHR

- Abnormal baseline rate
  - a) Tachycardia (FHR > 160 bpm)
  - b) Bradycardia (FHR < 110 bpm)
- Changing FHR - If an increasing or decreasing FHR is detected over time, potential causes should be explored and interventions initiated before the absolute values of bradycardia or tachycardia are reached.
- Presence of decelerations, especially if the FHR is slow to recover after a contraction

SE: ANNOTATE: Choose one of the situations of abnormal FHR above and explain it.
SQ: ANNOTATE: Create a question about one of the situations of abnormal FHR.
BN: ANNOTATE: Create a note.

CLINICAL MANAGEMENT

The data from intermittent auscultation should always be interpreted in conjunction with the **total clinical picture**. Interpretation of the findings is dependent on the stage of labour, the maternal clinical condition and fetal health prior to labour.

1. Normal

- Continue intermittent auscultation as per protocol.
- Continue to promote maternal comfort and fetal oxygenation, and to provide supportive care

2. Abnormal

- Interpret the abnormal findings in conjunction with the **total clinical picture.**
- Perform further **assessments** to confirm findings and determine potential causes:
    - Auscultate FHR again following the next contraction to confirm abnormal FHR
    - Assess potential causes;
    - Check maternal pulse, BP & temperature;
    - Perform a vaginal exam as indicated

SE: ANNOTATE: Explain the assessments done after detecting abnormal fetal heart rate (FHR).
SQ: ANNOTATE: Create a question about the assessments done after detecting abnormal fetal heart rate (FHR).
BN: ANNOTATE: Create a note.

- **Intervene** in an attempt to eliminate or reduce the effects of the cause, and to promote four physiologic goals
    - improve uterine blood flow;
    - improve umbilical blood flow;
    - improve oxygenation;
    - decrease uterine activity
- Consider the use of **additional fetal health surveillance measures:** digital fetal scalp stimulation, EFM, and/or fetal scalp sampling (if available).
- If not present, notify the primary care provider and consider delivery if the problem does not resolve

Note: The decision regarding the appropriate time to switch from intermittent auscultation to EFM is determined in part by the severity of the abnormal findings and on clinical judgment.

IMPLEMENTING AUSCULTATION INTO PRACTICE

Strategies for research utilization and knowledge transfer are key to implementing change in practice. Fischbeck-Feinstein, Sprague & Trépanier (2008) suggest a number of ways to support the implementation of IA in perinatal practice.

STRATEGIES TO PROMOTE IMPLEMENTATION OF IA

- Foster a culture of normal birth
- Develop unit policies to support IA
- Avoid EFM admission heart rate tests (admission strips) unless indicated
- Require documentation of an indication when EFM is used
- Educate all staff and physicians on IA techniques
- Recruit champions to model practice
- Work collaboratively with midwives
- Promote a work environment that encourages nurse presence and support
- Implement IA in stages
- Celebrate and share success
- Conduct regular audits and share outcome data
- User-friendly documentation tools
- Involve women in decision making about FHS

SE: ANNOTATE: Create a question about the strategies to promote implementation of IA.
SQ: ANNOTATE: Create a question about the strategies to promote implementation of IA.
BN: ANNOTATE: Create a note.

Appendix E.

Interpolated Task

1. How would you describe your general technology skills?

○ a) Not at all familiar
○ b) Slightly familiar
○ c) Somewhat familiar
○ d) Moderately familiar
○ e) Extremely familiar

2. How would you describe your familiarity with online learning environments?

○ a) Not at all familiar
○ b) Slightly familiar
○ c) Somewhat familiar
○ d) Moderately familiar
○ e) Extremely familiar

3. How many for-credit courses have you taken in an online-only format?

4. Any comments on the animations used in this chapter? Did you find them helpful?

5. Please provide any comments on your experience using nStudy software in this study

6. Please provide any comments or feedback on the presentation of the FHS Manual content, such as areas for improvement in the presentation, etc.

Appendix F.

Scoring Rubric for Posttest Responses

1. Describe the recommended frequency of auscultation during the active second stage of labour.

   Every 5 minutes

   Out of 1 mark. Targets **incidental** learning.

2. What Fetal Heart Rate characteristics can be assessed by intermittent auscultation?
   - Baseline fetal heart rate (1)
   - Rhythm (1)
   - Presence of accelerations (1)
   - Presence of decelerations (1)

   Out of 4 marks. Targets **incidental** learning.

3. What factors affect how you should interpret findings from intermittent auscultation?
   - stage of labour (1)
   - maternal clinical condition (1)
   - fetal health prior to labour (1)

   If they wrote: "Total clinical picture" in absence of the above = 1 mark total

   Out of 3 marks. Targets **incidental** learning.

4. For each of the following four cases, indicate whether Intermittent Auscultation or Electronic Fetal Monitoring would be preferred:

   A) After 42 weeks gestation (post-term pregnancy): EFM (1)

   B) After 41+3 weeks gestation, with normal non-stress tests (NSTs) and amniotic fluid volume: IA (1)

C) Pre-term, below 36 weeks gestation: EFM (1)

D) Vaginal Birth After Cesarean: EFM (1)

Out of 4 marks. Targets **incidental** learning

5. What is the recommended frequency of auscultation after the initial epidural dose?

   Every 5 minutes (1) for 30 minutes (1)

   Out of 2 marks. Targets **intentional** learning.

6. List up to ten strategies to promote implementation of Intermittent Auscultation:

   - Foster a culture of normal birth
   - Develop unit policies to support IA
   - Avoid EFM admission heart rate tests (admission strips) unless indicated
   - Require documentation of an indication when EFM is used
   - Educate all staff and physicians on IA techniques
   - Recruit champions to model practice
   - Work collaboratively with midwives
   - Promote a work environment that encourages nurse presence and support
   - Implement IA in stages
   - Celebrate and share success
   - Conduct regular audits and share outcome data
   - User-friendly documentation tools
   - Involve women in decision making about FHS

   1 mark each. Out of 10 marks total. Targets **intentional** learning

7. What is the recommended frequency of auscultation during the latent phase if the woman is in hospital during this phase, and at low-risk?

   At time of assessment (1) and approximately every hour (1)

Out of 2 marks. Targets **intentional** learning

8. What are two examples of non-electronic auscultation devices?

   Fetoscope (1) or Pinard-type stethoscope, (1), or doppler (1)

   Out of 2 marks. Targets **intentional** learning.

9. What can't be assessed by Intermittent auscultation?

   Baseline variability. (1)

   Classification/type of deceleration. (1)

   Other visually identifiable FHR changes.(1)

   Out of 3 marks. Targets **intentional** learning.

10. Why is it important to perform Leopold's Maneuvers prior to auscultation?

    To identify where the fetal back is (1) in order to best listen to the fetal heart. (1)

    Out of 2 marks. Targets **intentional** learning.

11. When should Fetal Heart Rate be assessed?

    - initiation of labour-enhancing procedures (e.g. amniotomy)
    - administration of medications
    - transfer or discharge of patient
    - admission of patient
    - artificial or spontaneous rupture of membranes
    - vaginal examinations
    - abnormal uterine activity patterns (e.g. increased basal tone or tachysystole)
    - any untoward event during labour (e.g. maternal hypotension, bleeding)
    - administration or initiation of analgesia/anaesthesia.

    1 mark each. Out of: 9 marks. Targets **intentional** learning.

12. What are the four physiological goals that you should promote as you intervene to eliminate abnormal FHR?
    - improve uterine blood flow; (1)
    - improve umbilical / placental blood flow; (1)
    - improve oxygenation; (1)
    - decrease uterine activity (1)

    Out of 4 marks. Targets **incidental** learning.

13. Describe the fourth maneuver in Leopold's Maneuvers.

    Face the client's feet. (1) Place both hands on the lower abdomen, (1) and move the fingers of both hands gently down the sides of the uterus toward the pubis. (1) Note the cephalic prominence or brow on the opposite side of the fetal back. (1)

    Out of 4 marks. Targets **incidental** learning.

14. In what criteria should a practitioner be experienced for intermittent auscultation?

    Practitioners must be experienced in the technique of auscultation (1), the palpation of contractions (1), and the auditory recognition of pertinent FHR changes (1).

    Out of: 3 marks. Targets **intentional** learning.

15. Name all the characteristics that indicate an abnormal fetal heart rate.

    - Tachycardia (FHR > 160 bpm)
    - Bradycardia (FHR < 110 bpm)
    - Changing FHR
    - Presence of decelerations

    1 mark each. Total out of 4 marks. Targets **intentional** learning

16. When would you use EFM after insertion of an epidural?

    If maternal hypotension occurs (1) and is a persistent problem (1).

Out of 2 marks. Targets **intentional** learning

17. What should you do before you auscultate the fetal heart?

    Leopold's Maneuvers. (or: Palpate the maternal abdomen to identify fetal presentation and position).

    Out of 1 mark. Targets **incidental** learning.

18. Describe the recommended frequency of auscultation during the active first stage of labour.

    Every 15 – 30 minutes (1 mark) and before the onset of pushing in the second stage of labour (1).

    Out of 2 marks. Targets **incidental** learning.

Appendix G.

Participant Responses to Interpolated Task

In the table below, responses to the questions about general technology skills and familiarity with online learning environments correspond to the following response choices:

1 = a) Not at all familiar

2 = b) Slightly familiar

3 = c) Somewhat familiar

4 = d) Moderately familiar

5 = e) Extremely familiar

Table G1. Participants' professions, previous experience in fetal health surveillance, reported general technology skills, and familiarity with online learning environments

Condition	Profession	Number of months doing fetal health surveillance	How would you describe your general technology skills?	How would you describe your familiarity with online learning environments?
SQ	Registered Nurse	0	4	3
SQ	Registered Nurse	0	4	4
SQ	Registered Nurse	30	5	5
SQ	Registered Nurse	6	4	3
SQ	Registered Nurse	0	3	4

SQ	Registered Nurse	3	4	3
SQ	Registered Nurse	0	3	3
SQ	Registered Nurse	18	5	5
SQ	Registered Nurse	3	3	5
SQ	Medical Doctor	2	4	4
SQ	Medical Doctor	18	4	4
SQ	Midwife	75	4	4
SQ	Nursing Student	0	4	4
SQ	Nursing Student	0	3	3
SQ	Nursing Student	0	5	4
SQ	Nursing Student	0	5	5
SQ	Medical Student	0	1	4
SQ	Student Midwife	0	5	5
SQ	Student Midwife	0	4	2
SQ	Student Midwife	0	3	1
SQ	Student Midwife	6	2	3
SE	Registered Nurse	0	5	5
SE	Registered Nurse	0	4	3
SE	Registered Nurse	0	1	4

SE	Registered Nurse	0	3	2
SE	Registered Nurse	240	3	4
SE	Registered Nurse	0	5	4
SE	Registered Nurse	48	4	4
SE	Registered Nurse	0	3	4
SE	Midwife	396	3	3
SE	Midwife	15	2	1
SE	Medical Student	0	3	2
SE	Medical Student	0	4	4
SE	Student Midwife	0	3	3
SE	Student Midwife	0	4	4
SE	Student Midwife	0	4	4
SE	Student Midwife	0	4	4
SE	Student Midwife	4	2	4
BN	Registered Nurse	0	3	5
BN	Registered Nurse	9	4	4
BN	Medical Doctor	192	4	4
BN	Midwife	144	4	5
BN	Midwife	48	4	4

BN	Nursing Student	0	5	5
BN	Medical Student	0	4	3

Note: SQ = Self-questioning, SE = Self-explanation, BN = Blank note

Appendix H.

Participant Comments

Table H1 Individual participant comments on their experience using nStudy annotation software in the study, by condition

Condition	Comments
SQ	nStudy software is a helpful tool for online learning and note-taking.
SQ	I didn't like the templates. They were good for cuing you to form a question, but I had a few times where I could not write the question I wanted because it didn't fit one of the templates. I also found it extremely difficult to read the content of the chapter while having to make up questions in the middle.
SQ	I found it really slowed down my studying having to try to make up questions that fit into the parameters in the box.
SQ	this was not helpful. spent more time focusing on how to manipulate the content into awkwardly worded fill in the blanks than I did not actually reading and interpreting the content on the chapter. I skimmed the information in the chapter to find something that would fit the provided boxes as it was required to complete this section. This exercise completely detracted from learning the material in the chapter
SQ	[left blank]
SQ	creating questions impaired my study process, my groove of reading and creating notes, I didn't like the software, Courses online that I have enjoyed were the Desousa institute. I like completing multiple choice questions

SQ	It was frustrating, I didn't like it. I will not use it again.
SQ	no comment
SQ	[left blank]
SQ	I don't like that this software was only available on Chrome, not on Safari
SQ	It was easy and user friendly to use and navigate.
SQ	Not sure what nStudy was doing but I did not like that I couldn't see the question I typed in the box, couldn't remember which ones I've done or not or if it saved
SQ	It was tedious and time consuming and I dont think it was helpful to me
SQ	creating the questions was confusing. I wasn't clear on whether I was creating questions based on ACTUAL content questions that I had, or simply making up questions as if I were testing someone else. I had no idea what types of questions you were looking for. I did not like this part at all.
SQ	[left blank]
SQ	user friendly, clear instructions
SQ	When I accidentally highlighted the entire page in blue, I was not able to unhighlight the page. So I had to read the text in blue, which made reading slightly difficult.
SQ	the annotate function was new - lots of potential
SQ	I did not like the fact that I didnt have an option to create my own question totally from scratch but that I had to fill in the blanks

SQ	I found it very distracting while I was reading.
SQ	It is an interesting concept. As a student, I hope to get the answers from the questions I have formed. Despite not being able to see the questions I have formed, I believe I was able to make connections with the reading material and what I already know through forming the questions. At times, I found myself forming questions that are not related to the annotation instructions.
SE	nstudy is a learning experience for me.
SE	I have no experience before this course but I found it easy to use and annotate helped allow me to think while I studied.
SE	[left blank]
SE	Why can't I save and review my notes...seems pointless
SE	[left blank]
SE	Layout was not familar to me, not overly easy to follow for someone with zero experience with this. More visual pictures. Bullet points not easy to follow for someone new to this learning
SE	It was at first hard to figure out how to navigate through this course. I didnt find the instructions helpful.
SE	good experience, easy to understand
SE	I did not feel at all confident.
SE	I found the software inconvenient and therefore, didn't use it as a study tool throughout.
SE	It was clear once it was downloaded but it was weird to have the annotation of what you knew then all the information below it. It was

	too hard to not read the following before addressing the question. The question should be first on a separate slide and then the information should follow on a different page
SE	very impressive, and provides a very easy aid. worked well for me
SE	Took a bit of figuring out but not over hard to use.
SE	I loved it, I think nStudy is such a good idea and I think it really helped me to think more deeply about the course material and solidify it in my own mind.
SE	It wasn't bad, I don't know if I would personally use it all the time
SE	Had to stop and re-read things that I did not understand to better annotate
SE	I loved it, good use of note taking technology, easy to use.
NP	[left blank]
NP	The softward was fine. Someitmes it was hard to figure out what to put into the note
NP	I appreciate that there was some diversity included in the animations. I wonder how the use of the study notes we create will become part of one's learning....will you get a set of notes at the end of each chapter that you could print out?
NP	Like the ability to make notes in a chapter. Makes material being covered more relevant to your own experience and encourages clinical thinking
NP	Good experience

NP	nice to make notes, would ove to have access to them whenever I would want to make a brief revue of the study material
NP	I had an Issue with the highlighting once, but it was fine