Photocopy and Use Authorization

In presenting this dissertation in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my dissertation for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this dissertation for financial gain shall not be allowed without my written permission.

Signature ___________________________________

Date _______________________________________

Examining the Effect of Scenario-based e-learning and Feedback Types on Learning Outcomes and Motivation

by

Sacha Johnson

A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in the Department of

Organizational Learning and Performance

Idaho State University

Summer 2020

© 2020 Sacha Johnson

Committee Approval

To the Graduate Faculty:

The members of the committee appointed to examine the dissertation of SACHA JOHNSON find it satisfactory and recommend that it be accepted.

Dr. David Coffland,
Major Advisor

Dr. Dorothy Sammons-Lohse,
Committee Member

Dr. Dani Moffit,
Committee Member

Dr. William Rose,
Committee Member

Dr. Robert Tokle,
Graduate Faculty Representative

Human Subjects Committee Approval

Idaho State University
Office for Research - Research Outreach & Compliance
921 S. 8th Avenue, Stop 8046 • Pocatello, Idaho 83209-8046

August 5, 2019

Sacha Johnson
Instructnl Tech Res. Ctr ITRC
MS 8064

RE: regarding study number IRB-FY2019-267: EXAMINING THE EFFECT OF SCENARIO-BASED E-LEARNING AND FEEDBACK TYPES ON LEARNING OUTCOMES AND MOTIVATION

Dear Ms. Johnson:

I agree that this study qualifies as exempt from review under the following guideline: Category 1. Research, conducted in established or commonly accepted educational settings, that specifically involves normal educational practices that are not likely to adversely impact students' opportunity to learn required educational content or the assessment of educators who provide instruction. This includes most research on regular and special education instructional strategies, and research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.
This letter is your approval, please, keep this document in a safe place.

Notify the HSC of any adverse events. Serious, unexpected adverse events must be reported in writing within 10 business days.

You are granted permission to conduct your study effective immediately. The study is not subject to renewal.

Please note that any changes to the study as approved must be promptly reported and approved. Some changes may be approved by expedited review; others require full board review. Contact Tom Bailey (208-282-2179; fax 208-282-4723; email: humsubj@isu.edu) if you have any questions or require further information.

Sincerely,

Ralph Baergen, PhD, MPH, CIP
Human Subjects Chair

Phone: (208) 282-1336 • Fax: (208) 282-4723 • isu.edu/research
ISU is an Equal Opportunity Employer

Dedication

I would like to dedicate this dissertation to my entire network of family and friends. Though there isn't enough room to name all of you here, I hope you all know how much your love and support mean to me. My husband, Mike Johnson, has always been one of my biggest cheerleaders and this wouldn't have been possible without his support. I would also like to dedicate this work to my mom and dad (Laurie and Jim), in-laws (Amy and Dave), sisters (Megan and Alma), brothers (Chris and Josh), cousins (Shanna and Coty), aunts (Chris and Pam), and uncles (Mike and John) who have all provided support throughout this project. Last, but certainly not least, I'd like to dedicate this to my good friends (Shannon and Sandy). All of you played a huge part in lifting me up during stressful times and without that, I would never have been able to accomplish this task.

Acknowledgments

I would first like to thank my committee members for sharing their knowledge and expertise. Special thanks goes to my major advisor, Dr. David Coffland, for guiding me along this journey. I would also like to acknowledge Dr. Dorothy Sammons-Lohse, Dr. Dani Moffit, Dr. William Rose, and Dr. Robert Tokle for serving on my committee and providing valuable input and support.

I would like to thank Susan Galindo for her willingness to be part of this research. It has been a pleasure to work with such a forward thinker who strives to provide students the very best educational experience. I must also thank all of the MLS professionals (Dave, Kristine, Traci, Angela, MaryAnne, Anita, and Heidi) who provided their expertise for the design of the instruction and assessments.

Lastly, I would like to thank Dr. Ed Galindo, Katrina Rhett, Jennifer Abbruzzese, and Jeremy Curtis for their feedback and support. Your help and encouragement were vital for the completion of this project.

Table of Contents

List of Figures

List of Tables

Examining the Effect of Scenario-based E-learning and Feedback Types on Learning Outcomes and Motivation

Dissertation Abstract--Idaho State University (2020)

The goal of this study was to examine differences in student learning outcomes and motivation over time as a function of scenario-based e-learning (SBeL) and feedback types (elaborative and intrinsic) as well as the interaction between these two variables. Research Question One asked if there was a statistically significant main effect of scenarios, main effect of feedback type, or interaction between the two variables on learning outcomes over time. Research Question Two asked if there was a statistically significant main effect of scenarios, main effect of feedback type, or interaction between the two variables on participants' motivation over time. A researcher-developed instrument was used as a pretest, immediate posttest, and delayed posttest to gauge participants' knowledge and skill acquisition. Keller's (2010) Instructional Materials Motivation Survey (IMMS) was used to determine participants' motivation. The results from these measurements were analyzed using two 2×2 repeated measures analysis of variance (ANOVA) with between-group factors and random assignment. The results showed no statistically significant main effects of scenario level or feedback type over time and no statistically significant interaction between these variables on participants' knowledge or motivation. Therefore, there was not sufficient evidence to reject the null hypotheses for either research question. As could be expected as a result of receiving the instructional intervention, a statistically significant main effect of time was observed for participants' scores between the knowledge pretest and immediate knowledge posttest.

Interestingly, a statistically significant main effect of time was also observed for participants' scores between the immediate and delayed knowledge posttests suggesting longer-term gains in knowledge; however, mastery levels decreased from immediate to delayed posttests suggesting there were no gains in skill acquisition. A statistically significant main effect of time on the IMMS instrument was also observed. This could be interpreted as participants having lost motivation after receiving the instructional treatments, which could be due to a lack of confidence in their ability to succeed on the knowledge instrument. The results of this study contribute to research into online learning, SBeL, feedback, and motivation.

Key Words: online learning, scenario-based e-learning, feedback, motivation, ARCS model

CHAPTER I

Introduction

The non-traditional students of today have driven demand for distance education (Allen & Seaman, 2016; Bozorgmanesh, Sadighi, & Nazarpour, 2011; Gibson, 2008; Stonebraker, Robertshaw, & Moss, 2016). According to Bozorgmanesh et al. (2011), distance education meets specific educational needs and its flexibility allows learners to study at any time and from anywhere, which fits today's learners who may work, have families, or otherwise not have access to higher education (The World Bank, 2002). Distance education has morphed from correspondence courses into online courses (Bozorgmanesh et al., 2011), which are being used to supplement or replace traditional classroom lectures and laboratories (Brinson, 2015; Allen & Seaman, 2013; Carnevale, 2003; Herrington & Oliver, 1997; Herrington, Oliver, & Reeves, 2003; Kee, Matthews, & Perumalla, 2009; Meisner, Hoffman, & Turner, 2008; Toth, Morrow, & Ludvico, 2009; Tzu-Chien, 2005); however, even with the growing number of online courses, many students report a preference for in-person courses (Tichavsky, Hunt, Driscoll, & Jicha, 2015). Tichavsky et al. (2015) stated, "students with and without experience with online courses seem to view online courses as lacking in interaction" (p. 3). According to these authors, the students mentioned instructor interaction more than other types of interaction and 26% of the students cited immediate instructor feedback as one reason for their preference for face-to-face classes. Ertmer et al. (2007) also noted the importance of feedback, stating "While instructor feedback is often cited as the catalyst for student learning in online environments, lack of feedback is most often cited as the reason for withdrawing from online courses" (p. 80). More recently, Hart (2012) suggested isolation, decreased engagement, and lack of relevance are barriers to persistence in online courses. Based on the results of their study, Park and Choi (2009) suggested, "lower dropout rates can be achieved if online program developers or instructors find ways to enhance the relevance of the course" (p. 207).

Approximately 30 years ago, Brown, Collins, and Duguid (1989) claimed that meaning comes from the context of use and suggested immersing learners in authentic practice. Scenario-based learning (SBL) emerged from that movement of valuing contextual knowledge (Errington, 2011). Errington (2011) claimed SBL "may provide one useful means for getting students closer to the realities of their intended profession through the construction and deconstruction of authentic learning experiences" (p. 84).

Combining these two elements, SBL and e-learning, may offer "opportunities to gain experience in a safe and controlled manner" (Clark, 2013, p. 12). However, before spending considerable time and money investing in new training methods, products, or designs, such as scenario-based e-learning (SBeL), Clark (2013) recommended asking critical questions such as, "*Does it work?*" or "*Does it work better than a different method or approach?*" as well as "*For what kinds of outcomes and for what kinds of learners is scenario-based e-learning most effective?*" (p. 137). The goal of this study was to examine whether e-learning scenarios and feedback have an effect on learning outcomes and learner motivation for a specific procedure within the medical laboratory science field. This chapter introduces the background for this research study as well as its practical significance and significance within the field of instructional design.

Background and Context

The needs of today's students have driven widespread adoption of distance education (Allen & Seaman, 2016; Bozorgmanesh et al., 2011; Gibson, 2008; Stonebraker et al., 2016). What began as a way to meet the needs of rural and remote students has grown into a preferred learning method for many (Bozorgmanesh et al., 2011). Stonebraker et al. (2016) observed that undergraduate students are increasingly "taking some or all of their college courses online" (p. 176). According to Bozorgmanesh et al. (2011), "Distance Education is now undertaken by people with busy schedules, hectic lifestyles, special needs, and also those living in isolated areas" (p. 75). This is especially true for higher education institutions as indicated in a report by The World Bank (2002), "Tertiary education institutions will have to organize themselves to accommodate the learning and training needs of a more diverse clientele: working students, mature students, stay-at-home students, traveling students, part-time students, day students, night students, weekend students, and so on" (p. 29).

Distance education comes in numerous forms such as correspondence, e-learning, broadcast learning, teleconferencing, and online learning (Bozorgmanesh et al., 2011). Allen and Seaman (2016) suggested, "The notion of a 'distance' . . . changes from being geographically separated to one of time shifting" (p. 11). Rapid advances in digital technology such as Web 2.0, have transformed distance and online education, which has shattered "the traditional model of unidirectional instruction by supporting online multilateral exchanges of visuals, text, and audio within and outside of the learning community" (Clark & Mayer, 2011, p. 7).

Today's instructional formats can be classified as traditional, web facilitated, blended/hybrid, and online (Allen & Seaman, 2016). Allen and Seaman (2016) differentiated between these formats by the degree of content delivered online. These authors stated that traditional courses have no content delivered online, web-facilitated courses are typically face-to-face courses with some content, such as the syllabus and assignments, delivered online, and blended or hybrid courses have increasingly more content delivered online such that the course may have reduced face-to-face time. Lastly, Allen and Seaman defined online courses as those with 80% or more of the content delivered online. Allen and Seaman (2013) reported the number of institutions offering online courses and complete online programs increased from 35% in 2002 to 62% in 2012.

The rise of online instruction has been accompanied by increasing research examining its effectiveness, much of which continues to strengthen the case for its use (Brinson, 2015). According to Sener (2005), "significant energy has been put into establishing the 'equivalent' quality of online courses and programs relative to traditional ones, as evidenced by the compilation of hundreds of distance education studies that document the well-known 'no significant difference' phenomenon" (p. 1). Choules (2007) suggested, "Comparison with more 'traditional' teaching is unhelpful. The future is to look at what elearning can do and use it to its strength" (p. 216). More recently, Clark and Mayer (2011) stated, "From the plethora of media comparison research conducted over the past sixty years, we have learned that it's not the delivery medium, but rather the instructional methods that cause learning" (p. 14). Clark and Mayer asserted that although there are pitfalls, e-learning offers the promise of customized training, learning engagement, multimedia, and accelerated expertise.

Such online instruction is not only being used to replace traditional classroom lecture material from single courses to entire programs, but also to supplement or replace the traditional classroom laboratory experience (Allen & Seaman, 2013; Carnevale, 2003; Herrington & Oliver, 1997; Herrington et al., 2003; Kee et al., 2009; Meisner et al., 2008; Toth et al., 2009; Tzu-Chien, 2005). Similar to research examining the effectiveness of online lecture instruction, researchers have also found comparable learning outcomes between online and in-person laboratory instruction (Carnevale, 2003; Herrington & Oliver, 1997; Herrington et al., 2003; Kee et al., 2009; Meisner et al., 2008; Toth et al., 2009; Tzu-Chien, 2005).

It has been three decades since Brown et al. (1989) identified an apparent disconnect between knowledge and practical application in classroom learning. According to these authors, this disconnect is the difference between knowing what and knowing how. For example, Brown and his colleagues cited a 1987 study by Miller and Gildea in which children were taught vocabulary words using dictionary definitions and a few example sentences. Brown et al. explained, "learning words from abstract definitions and sentences taken out of the context of normal use . . . is slow and generally unsuccessful" (p. 32), as evidenced by the students' misuse of the vocabulary words. Six years after Brown et al. identified this disconnect, Choi and Hannafin (1995) agreed that, "Formal learning emphasizes abstract and systematic problem-solving strategies" (p. 54). Choi and Hannafin suggested that when knowledge is obtained in such an isolated and decontextualized way, the knowledge is inert, meaning the learner can recall it but cannot apply it in practice.

Brown et al. (1989) developed situated learning (or situated cognition) as a framework to address this disconnect. Collins (1991) stated, “Situated learning is the notion of learning knowledge and skills in contexts that reflect the way the knowledge will be useful in real life” (p. 2). Generally, learning from unrealistic classroom instruction tends to create abstract knowledge learners are unable to apply in practical situations (Choi & Hannafin, 1995). Learners should therefore be immersed within the culture or community of practice to learn the knowledge and how to use it as practitioners do (Brown et al., 1989). As Brown and his colleagues suggested, “authentic activity becomes a central component of learning” (p. 37). Using SBL is one way to immerse learners within realistic situations (Clark, 2013).

According to Errington (2010), SBL “refers to any educational approach that involves the use or dependence upon, scenarios to bring about desired learning intentions” (p. 2). Extending SBL to online education, Clark (2013) defined SBeL as “a preplanned guided inductive learning environment designed to accelerate expertise in which the learner assumes the role of an actor responding to a work-realistic assignment or challenge, which in turn responds to reflect the learner’s choices” (p. 5). Errington (2010) asserted that scenarios have the potential to provide rich practical experience that extends past conventional lectures and tutorials. Researchers have examined the use of scenarios as preparation for flipped learning in which lecture-based material is delivered online and then applied in person (Lehmann, Bosse, Simon, Nikendei, & Huwendiek, 2013). Lehmann et al. (2013) found that using SBeL as preparation for flipped learning allowed for more efficient use of laboratory time. Researchers have also studied the effects of scenarios on student motivation and learning (Breakey, Levin, Miller, & Hentges, 2008; Clark, 2016; Errington, 2010; Errington, 2011; Landrigan, 2010; Lee & Butler, 2003; Lim, Reiser, & Olina, 2009). Lim et al. (2009) found that learning transfer may be improved through the use of real-world context (whole task) versus a traditional step-by-step (part task) approach. Landrigan (2010) concluded that SBeL activities are most beneficial when accompanied by active teaching. Errington (2010) explained that the evidence regarding whether scenarios help students as would-be professionals is predominantly anecdotal. Based on the results of their study, Lee and Butler (2003) asserted, "not all authentic situations are appropriate for the development of students' understandings of scientific knowledge" (p. 27).

Clark (2016) also noted that SBeL is not appropriate for all learning outcomes or for all learners. Previously, Clark (2013) suggested that SBeL should be considered "in situations in which on-the-job experience is rare, dangerous, or impractical" (p. 13). Clark (2013) listed two major limitations of SBeL research, which were the low number of experiments and inconsistency in terminology and implementation of SBeL. Clark (2013) proposed, "We need more research comparisons of scenario-based with directive lesson designs in different problem domains and with different learners to make more precise recommendations regarding when and for whom scenario-based e-learning is the more effective approach" (p. 143). Clark (2013) posed the question of what aspects of SBeL designs, such as reflective feedback, add learning value.

Feedback is another factor found to affect learning outcomes (Clark, 2013; Hattie & Gan, 2011; Shute, 2008). Clark (2013) stated, "Detailed feedback . . . makes all the difference between an effective and an ineffective learning experience" (p. 103). Hattie and Gan (2011) suggested, "There is a preponderance of evidence that feedback is a powerful influence in the development of learning outcomes" (p. 249). Hattie and Gan reviewed 12 meta-analyses and, based on the results, placed feedback "among the top 10 influences on achievement" (p. 249) and also suggested certain types of feedback have an effect on learner confidence and motivation; however, the authors also found considerable variance among the effects of feedback. Shute (2008) conducted an extensive literature review and also suggested that the results of feedback research have been inconsistent and stated, "Care should be taken to know which interventions increase performance and under which conditions" (p. 170). Clark echoed this sentiment by stating, "From the research we have learned that feedback has huge potential to improve learning, but not all feedback is effective" (p. 104).

Shute (2008) stated that elaborated feedback provides more information regarding learner responses besides accuracy. Clark (2013) referred to this type of feedback as instructional feedback. According to Clark, instructional feedback indicates whether an answer is correct and "can explain the reasons for the outcomes and suggest alternative actions to consider" (p. 106). For the purposes of this study, the term elaborative will refer to this level of feedback.

The second type of feedback that will be used for this study is what Clark (2013) referred to as intrinsic feedback. Clark suggested that intrinsic feedback mirrors the real world by providing learners with the results or consequences of their actions. Keller (2010) suggested providing learners with feedback that indicates causes of mistakes and ways to correct the mistakes helps learners build confidence. Confidence is one component of Keller's Attention, Relevance, Confidence, and Satisfaction (ARCS) model approach for designing motivational instruction. According to Keller, the model "illustrates how motivation, which influences the amount of effort that a person will exert toward achieving a goal, combines with their knowledge and skills to influence their overall performance" (p. 6).

As mentioned previously, Clark (2013) suggested an SBeL approach could be considered when learning on-the-job is rare, dangerous, or impractical. As such, one area within health professions that may benefit from contextualized e-learning and appropriate use of feedback is medical laboratory science (MLS). Medical laboratory scientists are responsible for performing medical laboratory tests and reporting results to physicians to aid in diagnoses. Medical laboratory science programs include a laboratory component as well as clinical rotations to teach students necessary real-world skills. Medical laboratory scientists make a series of decisions, and immediate feedback is necessary to direct that decision process. Hofstein and Lunetta (2004) defined science laboratory activities as "learning experiences in which students interact with materials and/or with models to observe and understand the natural world" (p. 31). However, some medical laboratory tests cannot be conducted in person due to cost-prohibitive equipment and lack of specimens. In such cases, lecture-based teaching is the only available instruction. This lack of in-person laboratory instruction can be addressed through the use of online laboratory instruction, and Hofstein and Lunetta recommended that simulations could be used to "engage students in investigations that are too long or too slow, too dangerous, too expensive, or too time or material consuming to conduct in school laboratories" (p. 42). Certain skills-based applications taught within the MLS classroom laboratory include how to use a specific type of microscopic slide (the hemocytometer) to count cells from different human body fluids. Although many hospitals have automated cell counters for body fluids, which mostly eliminate the need for professionals to manually count cells (S. Galindo, personal communication, February 17, 2017; M. Manis, personal communication, September 6, 2017), this skill is still valuable in the event that such equipment produces questionable results (Kiechle, 2017). For example, Kiechle (2017) suggested manually counting platelets when platelets clump, which causes automated analyzers to present inaccurate results.

Clark (2013) also suggested using scenarios in instances where on-the-job training is impractical. One such instance is using human body fluids in classroom laboratories because of the difficulty procuring specimens and the possibility of exposure to infectious pathogens (S. Galindo, personal communication, February 17, 2017). In particular, sperm cell counts are difficult to teach within classroom laboratories and are also more challenging than counting other cell types. This is because sperm must be analyzed soon after collection and sperm cell shapes are not always consistent (S. Galindo, personal communication, February 17, 2017). While instructors at some institutions are able to procure animal sperm specimens from veterinary clinics, it is highly impractical to obtain human sperm samples for testing in traditional classroom laboratory settings (S. Galindo, personal communication, February 17, 2017). Additionally, if human semen samples could be obtained, it is difficult to obtain abnormal specimens and the specimens degenerate quickly, which limits the opportunity to analyze them (S. Galindo, personal communication, February 17, 2017). Although MLS students gain some practical experience through clinical rotations, it cannot be guaranteed that students will have the opportunity to count sperm cells during their rotations (S. Galindo, personal communication, February 17, 2017). The expansion of the MLS program at a medium-sized university in the Intermountain West to multiple campus locations, as well as to online settings, necessitates e-learning instruction for both lecture and laboratory components. Students who attend the program online are unable to attend in-person laboratory sessions. Furthermore, providing in-person laboratory instruction for MLS students is not possible in such cases where obtaining human specimens is impractical. Therefore, the opportunity for students to practice and build their skills and confidence for manually counting sperm cells is limited. Using SBeL instruction could provide an opportunity for students to learn and practice counting sperm cells and also guarantee that students experience counting abnormal cells (S. Galindo, personal communication, February 17, 2017). According to Rothmann and Reese (2007), "semen analysis is practically the last routine manual microscopic test in the laboratory" (p. 18) and is therefore disliked by many professionals. Over a decade later automation is fairly common; however, Rothmann and Reese also asserted that the test lacks popularity due to inadequate instruction and that professionals lack confidence in their results. Although semen analysis is not a standard test performed in all clinical laboratories, because manual cell counts are required in the event of discrepancies, abnormal findings, small sample sizes and equipment malfunctions, this is a valuable and necessary skill, which is transferable to manually counting various other cell types (Kiechle, 2017; S. Galindo, personal communication, February 17, 2017). Rothmann and Reese included inadequate training and lack of practice opportunities among some of the problems with semen analysis. The authors found professionals felt that, "Semen analysis is discussed, at most, for a few hours in medical-technology education and often is not included in clinical training" (p. 18). Additionally, because semen analysis may not be performed daily, "competency and speed are difficult to accumulate" (Rothmann & Reese, 2007, p. 18). The lack of instruction and practice for manually counting sperm cells coupled with the need to develop this skill and students' confidence could possibly be addressed using SBeL; however, it remains to be seen whether this skill can be taught online and whether scenarios or feedback have an effect on students' learning, confidence, or overall motivation.

Purpose of the Study

The comparison of a scenario-based to a non-scenario-based e-learning approach and the question of the value of appropriate feedback to build confidence and motivation were the catalysts for this research study. Based on the existing literature, it was thought that the context and relevance of scenarios could prompt students to think like medical laboratory scientists, which could improve learning, confidence, and motivation. For the purpose of this study, it was thought that presenting scenarios would connect the instruction to learners' chosen profession, which should increase the relevance of the instruction and therefore learners' motivation. Additionally, feedback has been shown to have an effect on learning (Clark, 2013), confidence, and motivation (Hattie & Gan, 2011) and was included in the study to determine whether e-learning scenarios with elaborative or intrinsic feedback are appropriate for teaching college students how to manually count sperm cells.

Research Design

The goal of this study was to evaluate mean differences in student learning outcomes and motivation as a function of scenarios and feedback type as well as the interaction between these two variables. Therefore, this study employed a pair of 2×2 repeated measures analysis of variance (ANOVA) with between group factors and random assignment. An advantage of the repeated measures design is the minimization of effects of individual differences (Gravetter & Wallnau, 2007). As shown in Table 1, two independent variables each with two levels, scenarios and feedback, were administered across four groups.

Table 1

Research Design

R	S	O_1	O_2	X_1	O_3	O_4	O_5	O_6
R	S	O_1	O_2	X_2	O_3	O_4	O_5	O_6
R	S	O_1	O_2	X_3	O_3	O_4	O_5	O_6
R	S	O_1	O_2	X_4	O_3	O_4	O_5	O_6

R Random assignment.
S Demographic survey.
O_1 Motivation pretest, Keller's (2010) Instructional Materials Motivation Survey.
O_2 Researcher-developed content pretest.
X_1 Non-SBeL module with elaborative feedback.
X_2 Non-SBeL module with intrinsic feedback.
X_3 SBeL module with elaborative feedback.
X_4 SBeL module with intrinsic feedback.
O_3 Immediate motivation posttest, Keller's Instructional Materials Motivation Survey.
O_4 Immediate researcher-developed content posttest.
O_5 Delayed motivation posttest, Keller's Instructional Materials Motivation Survey.
O_6 Delayed researcher-developed content posttest.

Participants completed Keller's (2010) *Instructional Materials Motivation Survey* (IMMS) as a pre-survey, immediate post-survey, and delayed post-survey to assess their motivation levels. Keller's IMMS consists of 36 items with Likert scale responses 1-5 (not true to very true). Participants also completed a content and performance pretest, immediate posttest, and delayed posttest to assess their ability to manually count sperm cells. The content and performance instrument was also designed to assess the instructional objectives outlined in Chapter 3 (page 96). The researcher-developed knowledge instruments consisted of true-false, drag-and-drop, and open-ended (numerical) questions. The pretest contained one absolute sperm cell count question and the immediate and delayed posttests included three absolute sperm cell count questions. Learner demographics were also collected and reported to provide a complete description of the sample, but were not used for anything other than descriptive statistics. The demographic survey collected participants' genders, ages, program type, and previous experience with the content and online delivery method. Participants' program type was collected on the demographic survey and reported because, depending on participation from outside institutions, it was possible the sample could include students in Medical Laboratory Science (MLS) and Medical Laboratory Technologist (MLT) programs. The major difference between MLS and MLT degrees is that MLS is a four-year program while MLT is a two-year program, which results in the requirement of different certification exams for each level. According to the Medical Technology Schools website, this is "a recognition of the differences in skill set" (n.d., para. 2). For example, the advanced training of medical laboratory scientists allows them to analyze findings, verify lab results, and communicate result information to other medical professionals. Furthermore, "This advanced education may be the reason why the MLS is given oversight and responsibility of the MLT in a laboratory setting" (Medical Technology Schools, n.d., para. 2).

During the first week of the intervention, each participant was randomly assigned to one of the four treatment groups, (a) no scenario with elaborative feedback (sE), (b) no scenario with intrinsic feedback (sI), (c) a scenario with elaborative feedback (SE), and (d) a scenario with intrinsic feedback (SI). All four treatments were designed to be effective to ensure successful learning for all participants regardless of treatment group, but the treatments varied based on the independent variables. As mentioned previously, SBeL refers to learners participating as an actor in a work-realistic challenge (Clark, 2013). Therefore, the SBeL treatments (SE and SI) for the current study framed the instruction in the context of working within a laboratory setting. For example, the first page of the instruction started, “Congratulations! You've just landed your first dream job in a medical laboratory. One of the first tests that have come in is a semen analysis.” Participants were then presented with scenarios where couples were unable to conceive as part of the instruction. The scenarios were designed to meet the relevance aspect of Keller’s (2010) ARCS model in particular. According to Keller, relevance is related to learners’ perceived usefulness of the instruction. Learners are more likely to pursue a goal if it is perceived to be useful. Keller stated, “a sense of relevance occurs when the content to be learned is perceived to be useful to one’s work . . . “ (p. 99). Although not all participants were from the MLS program, it was thought that including these factors in the instructional design might increase the perceived usefulness for all learners.

Alternatively, participants in the non-SBeL groups (sE and sI) received the same instruction, but without the scenarios. For example, the non-SBeL instruction started, “This lesson is designed to show you how to manually count sperm cells using a standard hemocytometer.” Participants in the non-SBeL groups were presented with the instructional objectives and walked through the steps to manually count sperm cells and calculate an absolute cell count without the context of working in a medical laboratory.

After group assignments in week 1, participants completed the demographic survey as well as the motivation and performance pretests. Keller's (2010) IMMS was used to ascertain whether the treatment conditions had an effect on learner motivation over time. During week four, participants completed the instructional intervention followed by Keller's IMMS to measure their motivation immediately after the instruction. Participants completed the immediate IMMS posttest prior to taking the immediate performance posttest, in order to avoid any effects of performance on motivation. During week seven, participants completed delayed motivation and content posttests. The mean difference between treatments over time was examined using the group results on the content pretest, immediate posttest, and delayed posttest.

Research Questions

To examine the impact of scenarios and feedback on student learning and motivation, the following research questions and sub-questions were addressed:

1. Does the use of SBeL and feedback type have a significant effect over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest?
    a. Does the use of SBeL have a significant main effect over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest?

b. Does feedback type have a significant main effect over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest?

c. Is there a significant interaction effect of the combination of SBeL and feedback type over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest?

2. Does the use of SBeL and feedback type have a significant effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's (2010) IMMS pretest, immediate posttest, and delayed posttest?

   a. Does the use of SBeL have a significant main effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's IMMS pretest, immediate posttest, and delayed posttest?

   b. Does feedback type have a significant main effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's IMMS pretest, immediate posttest, and delayed posttest?

   c. Is there a significant interaction effect of the combination of SBeL and feedback type over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's IMMS pretest, immediate posttest, and delayed posttest?

Definitions of Terms

Cognitive load: According to Clark (2014), cognitive load refers to, "The amount of mental work imposed on working memory" (p. 874). Cognitive load has three sub-categories:

- Extraneous cognitive load: Refers to the burden imposed on working memory that Mayer (2001) asserted, "depends on the way the instructional message is designed – that is, on the way material is organized and presented" (p. 50). Renkl (2014) clarified that extraneous load "is unnecessary for the achievement of learning goals" (p. 407).
- Germane cognitive load: According to Moreno (2004), "Germane cognitive load promotes learning by having students devote more cognitive resources to tasks that are relevant to schema acquisition through increased effort or motivation" (p. 101).
- Intrinsic cognitive load: The burden imposed on working memory that Mayer (2001) stated, "depends on the inherent difficulty of the material – how many elements there are and how they interact with each other" (p. 50).

Corrective feedback: According to Johnson and Priest (2014), corrective feedback is, "A feedback message that merely informs the learner that his or her response was correct or incorrect" (p. 460).

Elaborative (instructional) feedback: Providing "information about the accuracy of a solution, along with an explanation" (Clark, 2013, p. 107). For the proposed study, elaborative feedback will be used to indicate whether learner responses are correct and provide additional information regarding why the response was correct or incorrect. An example of elaborative feedback could be informing the learner that the sperm cell count was not correct and suggesting actions to rectify the error.

Instructional Materials Motivation Survey (IMMS): A self-report instrument "that can be used to estimate learners' motivational attitudes in the context of virtually any delivery system" (Keller, 2010, p.277). For the purposes of this study, Keller's (2010) IMMS will be used to measure learner motivation before, immediately after, and two weeks after an e-learning module.

Intrinsic (consequential) feedback: Delivering the results or consequences of learner actions or responses (Clark, 2013, p. 107). Regarding this study, intrinsic feedback will be used to show learners how their responses relate to the scenario of reporting manual sperm count results. An example of intrinsic feedback would be informing the learner that the sperm cell count was not correct by describing the consequences such as a couple unnecessarily spending thousands of dollars to conceive.

Medical Laboratory Science (MLS): This abbreviation refers to a four-year baccalaureate program for medical laboratory scientists.

Medical Laboratory Technician (MLT): This abbreviation refers to a two-year technical training program for medical laboratory technicians.

Cognitive processing: Clark and Mayer (2011) defined cognitive processing as, "attending to relevant material, organizing the material into a coherent structure, and integrating it with what [learners] already know” (p. 35). Cognitive processing can be divided into three categories:

- Essential processing: Mayer and Fiorella (2014) asserted that essential processing is, “Cognitive processing – such as selecting relevant words and images, organizing selected words and images, and integrating – required to make sense out of the essential material” (p. 309).
- Extraneous processing: According to Clark and Mayer (2011) extraneous processing is, “cognitive processing that does not support the instructional objective and is created by poor instructional layout” (p. 37). Similarly, Renkl (2014) defined extraneous processing as, “working memory processes that do not contribute to the acquisition of the ‘targeted’ knowledge” (p. 407).
- Generative processing: Clark and Mayer (2011) defined generative processing as, “cognitive processing aimed at deeper understanding of the core material (consisting mainly of organizing and integrating)” (p. 37). Renkl (2014) suggested generative processing is, “working memory processes that contribute to the acquisition of ‘targeted’ knowledge” (p. 407). Mayer (2014c) suggested generative processing “involves making sense of the material by reorganizing it into a coherent structure and integrating it with relevant prior knowledge” (p. 61). Mayer (2014a) defined generative processing as, “Cognitive processing during learning that is aimed at making sense of the essential material in the lesson and is caused by the learner’s motivation to exert effort” (p. 68).

Motivation: According to Keller (2010), "Motivation is generally defined as that which explains the *direction* and *magnitude* of behavior…it explains what *goals* people choose to pursue and how actively or intensely they pursue them" (p. 4, emphasis in original).

Scenario-based learning (SBL): Refers to "any educational approach that involves the use or dependence upon, scenarios to bring about desired learning intentions" (Errington, 2010, p. 2). "Scenarios may constitute . . . an incident within a professional setting" (Errington, 2010, p. 2). For the purposes of this study, SBL will refer to framing the instructional module within the context of receiving an order for a manual sperm count, performing the count, and reporting the results.

Scenario-based e-learning (SBeL): Refers to "a preplanned guided inductive learning environment designed to accelerate expertise in which the learner assumes the role of an actor responding to a work-realistic assignment or challenge, which in turn responds to reflect the learner's choices" (Clark, 2013, p. 5). Because the SBL for this study was delivered online, it was considered SBeL.

Schema: Rumelhart and Ortony (1977) defined schemas as, "data structures for representing the generic concepts stored in memory" (p. 101). Sweller (1999) asserted, "A schema is defined as a cognitive construct that permits us to treat multiple elements of information as a single element categorized according to the manner in which it will be used" (p. 10). Sweller explained this concept stating that although trees have unique appearances, "a tree schema, held in long-term memory, allows us to ignore most of the enormous perceptual detail of any tree that we see and instead, consider one tree in the same manner as we might consider any other tree" (p. 10).

Situated learning theory (SLT): According to David (2007), SLT is based on the premise that learning is "situated within authentic activity, context, and culture" (para. 1).

Delimitations

According to Bracht and Glass (1968), the external validity of experiments can be affected by twelve factors, which are broadly categorized under population validity and ecological validity. Population validity factors are those that deal with whether experimental results can be generalized to the broader population. These factors are the comparison of the accessible population to the target population and the interaction of participant characteristics to the treatment. Both of these factors had the potential to impact the results of the current study and are discussed next.

Experimentally accessible versus target population. The accessible population refers to those participants who are available to the researcher while the target population refers to the larger population of interest. In the case of the current study, the accessible population was the students at MLS and MLT programs from participating institutions and the larger population refers to all MLS and MLT students. It is not practical to include all possible MLS and MLT students across the nation and therefore the findings of this study may not generalize to the broader population. Although faculty at institutions across the United States expressed interest in participating in this study, the participants were not selected to represent the national population; therefore, the results may not be representative. Additionally, because a limited number of participants from MLS and MLT programs were obtained, students from other programs at the researcher's institution as well as instructional designers were also included in the sample. Furthermore, because the study materials were delivered online, students without access to the Internet were not able to participate, which created an opportunity for bias and also limited generalizability.

Interactions of personological variables and treatment effects. Personological variables refer to participant characteristics. Not all individuals are the same and therefore it is difficult to generalize results to a larger population without a representative sample. In an effort to obtain a representative sample, the researcher attempted to include participants from additional institutions, which could have expanded the generalizability of any observed treatment effects. Participation from outside the MLS program could also expand generalizability to a more general population. Additionally, this study employed random assignment, which should have balanced participant characteristics between the groups and minimized this effect.

In addition to population validity factors, Bracht and Glass (1968) outlined ten ecological validity factors, which deal with the external environment of the experiment and under which conditions similar results can be expected. These factors include explicit description of the independent variable, multiple treatment interference, Hawthorne effect, novelty and disruption effects, experimenter effect, pretest sensitization, posttest sensitization, interaction of history and treatment effects, dependent variable measurement, and the interaction of measurement timing and treatment effects. Because the current study employed a single treatment, the threat of multiple-treatment interference was eliminated. Each of the remaining factors that may have had an effect on the results of the current study are addressed below.

Describing the independent variable explicitly. For results to be replicated, the treatment and experimental design must be explicit. Every effort was made to outline the treatment and setting for this study so that it may be replicated with other participants and in other disciplines.

Hawthorne effect. A Hawthorne effect can occur when participants are aware of their involvement in an experiment, which subsequently affects their behavior and performance. To avoid this threat, the researcher informed participants of the study to obtain informed consent; however, participants were not informed of the particulars of the study prior to receiving the instructional treatment.

Novelty and disruption effects. Participants may perform differently due to a treatment's newness or its disruption. With the widespread use of the Internet and computer-assisted instruction, it was thought that participants would likely have previous experience with the online delivery method for the treatment; however, prior experience data was collected as part of the demographic survey. The results were reported to determine whether prior experience appeared to have an effect on the study results.

Experimenter effect. The person administering a study can cause unintended participant behavior. The treatment for this study was delivered online using a learning management system (LMS), which should have eliminated the possibility of experimenter effects. Although there were outside participants, no other MLS or MLT institutions participated in this study; therefore, the researcher delivered all materials using an LMS at the researcher's institution, which helped to ensure the treatment and control conditions were the same across participants.

Pretest sensitization. This threat refers to the possibility of familiarizing the participants to the treatment by administering a pretest. For the current study, participants were taught to manually count sperm cells. This procedure requires the use of a hemocytometer, which is a specific microscope slide that is also used for counting other cell types. It was hoped that the time delay of two weeks between the pretest and immediate posttest would help reduce this threat. Additionally, the pretest only included one absolute cell count question to avoid the pretest serving as instruction.

Posttest sensitization. Participants may not fully understand the content during a treatment, but be able to understand it upon seeing the information on a posttest. In this case it appears the learning occurred due to the treatment; however, it actually occurred due to the posttest. For the purposes of this study, the format of the practice opportunities provided in the instructional intervention were identical to the immediate and delayed posttests, but the assessment samples were each unique, which resulted in unique sperm cell counts for each item on the assessments. In addition, the posttests did not include any feedback elements. Furthermore, in an attempt to minimize this threat, there was a delay of two weeks between the immediate and delayed posttests.

Interaction of history and treatment effects. Student morale or other concurrent event during treatment duration could affect study results. Therefore, during the duration of this study any such occurrences were noted and taken into account when interpreting the results. Participants from outside the researcher's institution included instructional designers and faculty who would not have been taught cell counts during the study duration, which should have minimized outside influence on participants' ability to count sperm cells.

All iterations of the study were conducted over seven weeks, which was a relatively short time for substantial historical events to occur; however, during the last two iterations, a stay-at-home order was issued in the researcher's state due to COVID-

19. The threat of COVID-19 may have had a negative effect on participants due to unforeseen circumstances such as being unable to access necessary resources for completing the online research study or decreasing overall motivation to participate in the study.

Measurement of the dependent variable. The measurement used to assess the outcome of a study can limit the generalizability of a study's results. The knowledge instruments used for this study were developed using the expertise of faculty and working professionals to ensure face and content validity. The instruments were specific to this study in order to assess students' knowledge and practical skills for manually counting sperm cells using a hemocytometer. As such, other forms of assessment such as essay-type questions may not yield similar results.

Interaction of time of measurement and treatment effects. Participants may perform differently on assessments over time. This could be due to lack of retention or participants' attitude on assessment days. According to Bracht and Glass (1968), "An experimental design which includes the measurement of the dependent variables at several points in time will increase the ecological validity of the results" (p. 466). The current study included an immediate posttest and a delayed posttest to determine whether any observed treatment effects were sustained. Although there were only two weeks between the immediate and delayed posttests, because the course at the researcher's institution was an eight-week course, 25% of the course content was covered in those two weeks. Both synovial and serous fluids were covered during those two weeks at the researcher's institution. It was hoped to gather topic schedules from outside institutions; however, no outside MLS or MLT programs participated.

Limitations

Campbell and Stanley (1963) and Cook and Campbell (1979) outlined twelve factors that can influence an experiment's internal validity, which refers to how confidently observed effects can be attributed to the treatment rather than to confounding variables. These factors include history, maturation, testing, instrumentation, statistical regression, differential selection, experimental mortality, selection-maturation interaction, experimental treatment diffusion, compensatory equalization of treatment, compensatory rivalry by the control group, and resentful demoralization of the control group. These factors could create the appearance of a treatment effect, when in fact the observed effect was not due to the treatment itself, but rather due to one or more of these extraneous factors. Therefore, these factors must be controlled in order to strengthen the internal validity of the study. Because the participants for this study were adults, there should have been little effect of maturation on participant results and it was not considered as a possible influence on the study results. The study used random assignment to determine the four treatment groups, which should also have minimized the potential for a selection-maturation interaction. Random assignment should also have limited the threat of compensatory rivalry by the control group and resentful demoralization of the control group. The remaining factors may have impacted the results of the current study and are therefore discussed individually with regard to the researcher's attempt to minimize the impact of each.

History. This threat pertains to any activities that happen between measurements. If an event occurs to a large enough portion of the sample, it could have an effect on the results (Campbell & Stanley, 1963). Such an event could be participants going beyond the treatment instruction to learn the content from other resources such as tutors or other textbooks. Each iteration of this study took place over the course of seven weeks from pretest to delayed posttest. In that time, more self-directed participants may have taken it upon themselves to further research the content. However, using random assignment should have resulted in self-directed learners being equally dispersed among treatment groups.

Other cell count instruction could have had an effect on the results of this study; therefore, instructors were asked not to teach other cell counts during the study. Any cell count instruction, whether semen or other cell types, between the immediate and delayed posttests could have had an effect on the delayed posttest results. Randomly assigning participants to the treatment groups decreased the likelihood of a systematic set of errors due to participants receiving or seeking out additional cell count instruction. Therefore, random assignment made it unlikely that any outside instruction would make a difference in participants' performance across the treatment groups. Further, this risk would be most pertinent to participants in MLS or MLT programs because it is unlikely that university students in other programs would receive specific cell count instruction without being enrolled in such a program.

Additionally, the study was designed with two weeks between the pretest and immediate posttest and two weeks between the immediate and delayed posttests. These periods of time between measurements increased the likelihood of other events having an effect on the results; however, the relatively short study duration helped to minimize the possibility of history effects.

Testing. Administering more than one assessment during an experiment can threaten the validity of the results (Campbell & Stanley, 1963). By exposing participants to the content with multiple tests, they become sensitized to it. To avoid this threat, the pretest and two posttests for this study were similar, but had unique cell counts. Additionally, the interruption created by the delay between the pretest, immediate posttest, and delayed posttest could have helped to reset participants' familiarity with the assessment.

Instrumentation. Variability in measuring the dependent variable can have an effect on the accuracy of a study's results. Multiple observers or subjective scoring can create differences in measurement. The ability for students to manually count sperm cells was assessed using online knowledge and skills assessments, which should have eliminated the potential for subjective measurement error. Additionally, to assess student motivation, an established assessment, Keller's (2010) IMMS, was used. The IMMS is a well-accepted measurement with established validity and reliability estimates. The content validity of the knowledge/skills instrument was determined by a panel of experts and will be discussed in Chapter 3. In addition, random assignment should have served to evenly distribute any characteristics of the sample that might have been favored by a particular element of the researcher-developed instrument.

Statistical regression. A treatment may appear to have caused an effect in instances where a test group is chosen due to extreme scores. Participants in that group may appear to perform better upon retesting; however, this is due to the tendency of the participant scores to regress toward the mean of the population. Participants in this study were not selected based on extremely high or low scores; however, there was still a possibility that students could initially score well, but their performance on the delayed posttest could regress toward the mean. Therefore, participants in this study were randomly assigned to groups prior to taking the pretest in an attempt to evenly distribute extreme scores on the assessments. Essentially, if participant scores spiked high or low on any of the instruments, statistical regression toward the mean could have occurred; however, random assignment should have reduced the likelihood of that occurring.

Experimental mortality. The loss of participants from one treatment group introduces the possibility that the group will no longer be representative of the sample. For the proposed study, this threat should have been addressed by the fact that students in the program had chosen this profession and were therefore vested in the program. These students were likely motivated to complete the required course work, which included the treatment for this study. The participants were taking the required course online and the treatment was delivered online in a way similar to the regular course materials. Therefore, it was hoped that students would not drop out of the research due to the treatment. The researcher offered Amazon gift card incentives to boost participation and persistence especially for non-MLS participants. A subset of the non-MLS participants were also offered extra credit in a university course for completing the study materials. It was hoped these efforts would reduce drop-out rates for non-MLS participants.

Experimental treatment diffusion. The possibility of participants in one group being influenced by members of the other group could affect the outcomes of a research study. The participants at the researcher's institution were students who took an online course, which could have helped keep the treatment groups separate; however, some of these students also attended classes on campus and could therefore have discussed the treatments with each other. The treatments were delivered in an online course that required users to login with a username and password. This level of security should have prevented students in one treatment group from experiencing the other treatment without the aid of another student.

Compensatory equalization of treatment. This threat refers to the possibility of compensating members of one treatment group, which could have an effect on their performance. None of the participants were treated differently based on their treatment group. Participants during the first iteration were enrolled in the Urinalysis and Body Fluids course at the researcher's institution. The course instructor was unaware of the participants' assigned treatments and was only informed whether the participants completed all of the study materials; therefore, the instructor did not compensate members of any particular group. There were participants in the second iteration who received extra credit for participating; however, the instructor for that course was also unaware which treatments the participants were assigned and therefore could not compensate members of one group over another. Participants in the third and fourth iterations did not receive course credit and no faculty members were made aware of their participation in the study.

Significance of the Study

Although SBL has been around for forty years (Clark, 2013) and there is a significant amount of research regarding its use in medical education, the focus of the research within the medical field has been on doctor and nurse education rather than MLS professionals. According to Clark (2013), "Two major limitations regarding various forms of scenario-based e-learning include a rather low number of experiments conducted to date and inconsistency in terminology and implementation of scenario-based e-learning" (p. 139). It is still unclear which circumstances and for which learners SBeL is an effective instructional tool; however, even though Clark stated high-end technology and complex skills are not required to create SBeL, if its use does not facilitate higher learning outcomes or student motivation, it would not be cost effective to create such instruction. The results of the current study could help to expand on the number of experiments as well as best practices for the implementation of SBeL within MLS education. Clark supports the use of SBeL for accelerating expertise in those circumstances where real-life practice is impractical, which is the case for manual sperm cell counts.

While most hospitals have automated cell counters for body fluids, MLS and MLT professionals must be able to manually count different cell types in order to pass required proficiency tests and continue performing such tests if their laboratories perform these analyses (S. Galindo, personal communication, April 17, 2019). Manual cell counts are required in the event of discrepancies, abnormal findings, small sample sizes and equipment malfunctions, which means it is a valuable skill that is transferable to manually counting various other cell types (Kiechle, 2017; S. Galindo, personal communication, February 17, 2017). Currently, sperm cell counts are difficult if not impossible to teach in a traditional in-person laboratory due to the difficulty obtaining and preserving samples. Therefore, most students are required to learn the procedure strictly through lecture and text-based procedures. If found effective, the instructional intervention designed for this study could serve to educate MLS and MLT students before entering the workforce and provide a valuable job aid for working professionals in the event of equipment malfunction, which could eliminate the need to delay performing tests while waiting for equipment to be serviced.

CHAPTER II

Literature Review

The purpose of this study was to examine the efficacy of scenario-based e-learning (SBeL) for teaching manual cell counts and the effect of two different feedback types on learner performance and motivation. Therefore, the review of the literature starts with an overview of distance and online learning as well as how these formats have given rise to more complex, multimedia e-learning. Next, a theoretical framework consisting of schema theory, a cognitive theory of multimedia learning (CTML), and a cognitive-affective theory of learning with multimedia (CATLM) is presented along with a discussion of the instructional design principles derived from CTML and CATLM. After this, situated learning theory (SLT) is discussed as a foundation for scenario-based learning (SBL) and SBeL. Following that, an overview of research examining feedback types delivery, and timing is presented. Finally, research on motivation and Keller's (2010) Attention, Relevance, Confidence, and Satisfaction (ARCS) model for motivational learning design are discussed.

Online Learning

Distance learning began in the United States as correspondence courses delivered via postal service and continued to evolve alongside technological innovations such as radio, television, and the Internet (Casey, 2008). Definitions of distance learning, "include technology as an intrinsic quality" (Casey, 2008, p. 46) and distance learning is typically characterized by the spatial and sometimes temporal separation of the instructor from the students (Bozorgmanesh, Sadighi, & Nazarpour, 2011; Moore, Dickson-Deane, & Galyen, 2011). Casey (2008) asserted that since its advent, "distance learning programs have snowballed into online instructional delivery systems capable of granting doctoral degrees" (p. 45). Online learning can be difficult to define and is often used interchangeably with other terms such as web-based learning and e-learning (Moore et al., 2011). Online learning comes in various forms with varying degrees of course content delivered via the Internet (Allen & Seaman, 2016). Allen and Seaman (2016) designated face-to-face courses as both traditional and web-facilitated courses, which have between zero to 29% of the content delivered online. These authors defined blended or hybrid courses as those with between 30% and 80% of the content delivered online and online courses have 80% or more of the content delivered online typically with no face-to-face meetings; however, some online courses do require students attend virtual meetings at specific times using group chats or video conferencing, which highlights the difference between synchronous and asynchronous online learning (Bozorgmanesh et al., 2011).

Regarding the efficacy of online instruction, Sener (2005) stated, "significant energy has been put into establishing the 'equivalent' quality of online courses and programs relative to traditional ones, as evidenced by the compilation of hundreds of distance education studies that document the well-known 'no significant difference' phenomenon" (p. 1). Subsequently, online instruction is not only being used to replace traditional classroom lecture material from single courses to entire programs, but also to supplement or replace the traditional classroom laboratory experience (Kee, Matthews, & Perumalla, 2009; Meisner, Hoffman, & Turner, 2008; Toth, Morrow, & Ludvico, 2009).

According to Waldrop (2013), "In the sciences, the standard vehicle for teaching practical skills is the lab course" (p. 268) and he continued on to state that historically with distance education, students completed laboratory work using kits that allowed them to conduct experiments at home. However, Waldrop stated that, "education-technology researchers have been making substantial progress over the past decade" (p. 268) to bring practical laboratory experiences online. This study included instruction for a course at the researcher's institution where 100% of the course content is delivered online and in particular, a laboratory component of the course. Extensive research of online laboratory and skills-based instruction has been conducted (Brinson, 2015; Carpenter, Watson, Raffety, and Chabal, 2003; Meisner et al., 2008; Toth et al., 2009).

Online laboratories. In his meta-analysis, Brinson (2015) reviewed 56 studies since 2005 comparing non-traditional (remote or virtual) science laboratories (NTL) to traditional hands-on laboratories (TL). He coded for six different learning outcome categories: knowledge and understanding, inquiry skills, practical skills, perception, analytical skills, and social and scientific communication. Based on the results of his analysis, most of the studies showed higher learning outcomes for NTL for all learning outcomes; however, there were differences depending on the type of learning outcome category. Brinson stated, "Studies supporting higher achievement in NTL seem to place a lot of emphasis on content knowledge and understanding . . . whereas studies supporting higher achievement in TL seemed to rely heavily upon qualitative data related to student and/or instructor perception" (p. 228).

Meisner et al. (2008) compared improvement of physics understanding between students (N = 62) completing a traditional introductory-level physics course (n = 40) and those completing the course using *LabPhysics* software (n = 22). The authors found that students in both groups had improved scores, but that student taking the virtual course had greater improvement. All students in the traditional group and all but one student in the virtual group worked collaboratively, which could have affected the results. The authors also indicated that it is possible for faculty to be unaware of student understanding, particularly when students work in groups, which does not occur in the virtual environment. The authors concluded, “learning physics in a virtual environment, driven by exemplary pedagogy, may be a viable alternative to the standard method of instruction” (p. 100).

Toth et al. (2009) examined the effects of combining virtual and hands-on laboratory work with an inquiry learning approach (*N* = 39). The authors used the regular registration procedure to split the course into two laboratory sections. One section completed the virtual laboratory first while the other section completed the hands-on laboratory first. The authors found that both approaches significantly improved student knowledge, but did not find a significant difference between the groups. The authors concluded, “the quantitative analysis suggests to practitioners, as well as curriculum designers, that they may decide to choose any specific order of presentation without significant difference in effect on student-learning” (p. 342); however, the authors also found that the students more strongly favored completing the virtual work first and suggested future research examining student experiences during the inquiry process in both approaches.

As can be seen, technological advances have afforded the ability to deliver complex, multimedia instruction in the online environment. Mayer and Moreno (2003) define multimedia instruction as, “presenting words and pictures that are intended to foster learning” (p. 43). The words can be presented verbally (narration) or non-verbally (text) and the pictures can be static or dynamic, but according to the authors, the intent is to facilitate meaningful learning.

Theoretical Framework

Multimedia researchers have used learning theories such as the cognitive load theory to develop instructional design principles for creating quality multimedia instruction (Clark & Mayer, 2011; Moreno & Mayer, 2007). As Mayer and Moreno (2003) commented, "In pursuing our research on multimedia learning, we have repeatedly faced the challenge of cognitive load: Meaningful learning requires that the learner engage in substantial cognitive processing during learning, but the learner's capacity for cognitive processing is severely limited" (p.43).

Schema theory. Although schema theory is not central to the current study, it is necessary to understand schemas because of the role they play in cognitive load theory. Sweller, Van Merriënboer, and Paas (1998) stated that schemas are used for organizing and storing knowledge. Rumelhart and Ortony (1977) defined schemas as, "data structures for representing the generic concepts stored in memory" (p. 101). The authors explained schemas using an example:

> Although it oversimplifies the matter somewhat, it may be useful to think of a schema as analogous to a play with the internal structure of the schema corresponding to the script of the play. A schema is related to a particular instance of the concept that it represents much the same way that a play is related to a particular enactment of that play. (p. 101)

Rumelhart and Ortony (1977) identified four essential characteristics of schemas; schemas (a) have variables, (b) can be embedded within other schemas, (c) represent generic concepts with varying levels of abstraction, and d) represent knowledge rather than definitions. Additionally, Van Merriënboer and Kester (2014) explained that schema construction is the process of forming complex schemas by incorporating lower-level elements into higher-level schemas. The more complex schema can then be treated as a single element in working memory. The authors clarified, "a large number of elements for one person may be a single element for another, more experienced person who already has a cognitive schema available that incorporates the elements" (p. 113). In addition, Van Merriënboer and Kester suggested repeated use of schemas results in schema automation, which, according to Sweller et al. (1998), results in the ability to perform procedures "with minimal conscious effort" (p. 256). Van Merriënboer and Kester stated, "Expertise develops through two complementary processes, namely, schema construction and schema automation" (p. 113).

Cognitive load theory. According to Sweller et al. (1998), human cognitive architecture consists of "a limited working memory that deals with all conscious activities and an effectively unlimited long-term memory that can be used to store schemas of varying degrees of automaticity" (p. 258). Because schemas are treated as a single element, they can reduce working memory load. Additionally, demands on working memory are also somewhat reduced by the automation of schemas. Sweller (1994) asserted that the mental load of working, also called cognitive load, is influenced by a combination of intrinsic and extraneous factors. Intrinsic load is inherent in the content being presented and instructors have no control over this factor (Sweller, 1994). On the other hand, extraneous cognitive load as Sweller explained, "is artificial because it is imposed by instructional methods" (p. 307); therefore, proper instructional design not only reduces extraneous cognitive load, but also increases germane cognitive load, which is the effort required to build schemas. Sweller stated, "According to cognitive load theory, engaging in complex activities . . . that impose a heavy cognitive load and are irrelevant to schema acquisition will interfere with learning" (p. 301). Consequently, Sweller asserted that specific instructional design approaches can reduce extraneous cognitive load and direct learners' attention to fundamental features for building schemas such as goal-free and worked example strategies.

Cognitive theory of multimedia learning. Drawing from cognitive load, dual-coding, and constructivist learning theories, Mayer and Moreno (2002) posited a cognitive theory of multimedia learning (CTML). Mayer and Moreno gathered from dual-coding theory that people have separate information processing channels for verbal and visual material. From cognitive load theory, the authors derived that each of these channels has a limited amount of processing capacity. From constructivist learning theory, Mayer and Moreno determined that, "meaningful learning occurs when learners actively select relevant information, organize it into coherent representations, and integrate it with other knowledge" (p. 111).

As shown in Figure 1, Mayer (2001) used these assumptions to present a model of how people learn from multimedia presentations. The top and bottom rows of the model indicate there are separate channels for processing words and pictures (respectively) and knowledge construction occurs within the limited working memory. The arrows labeled selecting words, selecting images, organizing words, organizing images, and integrating represent the active cognitive processing required for meaningful learning. According to Clark and Mayer (2011), "The challenge for the learner is to carry out these processes within the constraints of severe limits on how much processing can occur in each channel of working memory at one time" (p. 37).

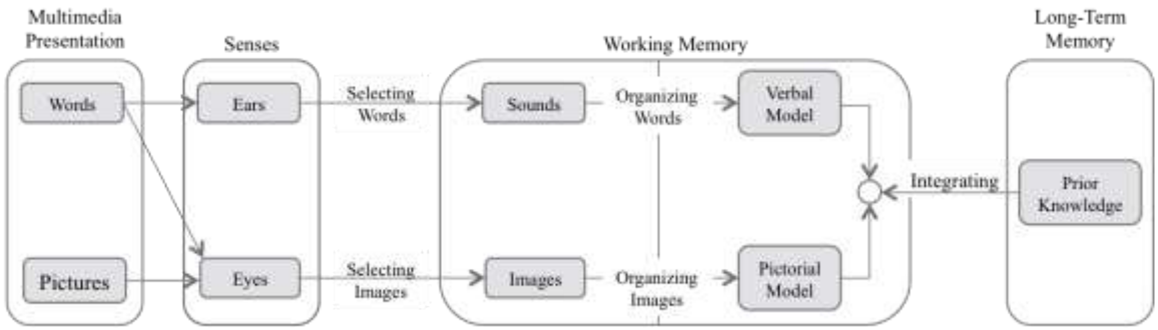


Figure 1. A cognitive theory of multimedia learning. Adapted from *Multimedia Learning* (1st Ed., p. 44), by R. E. Mayer, 2001, New York, NY: Cambridge University Press. Copyright 2001 by Cambridge University Press. Adapted with permission (see Appendix A).

Clark and Mayer (2011) stated that cognitive processing occurs within three categories. Extraneous processing equates to extraneous cognitive load within cognitive load theory and should be limited because it is a result of poor instructional design and does not support the learning outcomes. Essential processing, which corresponds to intrinsic cognitive load, is related to selecting relevant material and is due to the innate complexity of the material. On the other hand, generative processing "is created by the motivation of the learner to make sense of the material" (p. 37) and deals primarily with organizing and integrating material. According to Clark and Mayer, "The challenge for instructional professionals is that all three of these processes rely on the learner's cognitive capacity for processing information, which is quite limited" (p. 38). As such, Clark and Mayer outlined several multimedia principles to help manage these processing demands during learning based on the CTML.

Multimedia principles. Mayer (2001, 2014c) and Clark and Mayer (2011) presented the multimedia principles based on extensive research connecting learning outcomes to instruction with text and graphics. These include the multimedia, spatial contiguity, temporal contiguity, modality, redundancy, coherence, personalization, segmenting, and pretraining principles. Each principle will be discussed separately below.

Clark and Mayer's (2011) *multimedia principle* is based on the assumption that relevant cognitive processing, and therefore deeper learning, is more likely to occur when instruction uses graphics in addition to words (printed or narrated). The authors stated, "Multimedia presentations can encourage learners to engage in active learning by mentally representing the material in words and in pictures and by mentally making connections between the pictorial and verbal representations" (p. 71). Clark and Mayer explained that, based on the research, combining text and graphics appears to be more important for novice learners than more experienced learners because more experienced learners may have enough prior knowledge to learn just as deeply from text or graphics exclusively.

Building on the multimedia principle, the *spatial contiguity principle* outlines that text and corresponding graphics should be placed near each other. Mayer (2001) asserted that by placing corresponding words near to the graphics, "*learners do not have to use cognitive resources to visually search the page or screen and learners are more likely to be able to hold them both in working memory at the same time*" (p. 81, emphasis in original). Instead, by integrating words and graphics, learners can more easily hold them in working memory, which increases the ability to make more meaningful connections.

While spatial contiguity relates to space on a screen or page, the *temporal contiguity principle* relates to the timing of narration and graphics within multimedia instruction. As Mayer (2001) described, by presenting narration and corresponding graphics simultaneously, "*the learner is more likely to be able to hold mental representations of both in working memory at the same time*" (p. 96, emphasis in original). According to Clark and Mayer (2011), "when a lesson separates corresponding words and graphics, learners experience a heavier load on working memory" (p. 103), which is an extraneous load that can hinder deep learning.

For the *modality principle*, Clark and Mayer (2011) suggested that deeper learning occurs "when words explaining concurrent graphics are presented as speech rather than as on-screen text" (p. 124). This principle is particularly important when learners must simultaneously focus on a complex graphic or animation, the words are familiar, and the lesson is fast-paced. The authors explained that learners cannot fully attend to an animation or graphic if they must also attend to the printed words. Alternatively, to reduce the probability of overloading learners' visual channels, text should be narrated rather than printed on screen. Similar to the multimedia principle, research has shown a stronger modality effect for less skilled learners. Additionally, according to Clark and Mayer "If the material is easy for the learner or the learner has control over the pacing of the material, the modality principle becomes less important" (p. 128).

Further, the *redundancy principle* also applies to audio narration. In this principle, Clark and Mayer (2011) advocated against repeating on-screen text in the narration because this redundant on-screen text can overload learners' visual channels, which is termed the redundancy principle. In some cases, learners will spend more time comparing on-screen text with graphics, which can increase extraneous cognitive load. In other cases, learners may attend more to the on-screen text than the graphics, which can hinder learning. Nevertheless, there are certain instances when redundant on-screen text does not overload the visual processing channel. These boundary conditions that Clark and Mayer referred to include instances when no graphics are present, learners have ample time to process the text and graphics, the verbal material is complex or unfamiliar, or the text directs attention to relevant information. In general, Clark and Mayer suggested avoiding redundant on-screen text. However, the authors noted that such text "may be appropriate when there are no concurrent graphics, the text is unfamiliar to the learner, the printed words are unobtrusive, or you can use the printed words to signal where to look on the screen" (p. 146).

According to Clark and Mayer's (2011) *coherence principle*, "adding interesting but unnecessary material – including sounds, pictures, or words – to e-learning can harm the learning process by preventing the learner from processing the essential material" (p. 161). Extraneous audio competes for limited auditory channel resources while unnecessary words or graphics can be distracting to learners. Because the coherence principle research Clark and Mayer examined dealt primarily with novice learners, the authors explained the possibility of a stronger coherence effect for novice learners. The authors suggested future research examining the effect of the coherence principle on different types of learners. Essentially, with respect to the coherence principle, Clark and Mayer suggested, "the challenge for instructional professionals is to stimulate interest without adding extraneous material that distracts from the cognitive objective of the lesson" (p. 173).

Clark and Mayer's (2011) *personalization principle* deals with using conversational style in instruction especially when on-screen agents are used. The authors stated, “The psychological advantage of conversational style, pedagogical agents, and visible authors is to induce the learner to engage with the computer as a social conversational partner” (p. 180). Clark and Mayer suggested that a conversational style mirrors human conversation and can help learners engage with the computer.

Similarly, the *voice principle* pertains to social cues (Mayer, 2014b). Clark and Mayer (2011) asserted that learning is better when narration is spoken in a human rather than machine voice. According to Mayer (2014b), “A machine-synthesized voice . . . may not strongly convey the idea that someone is speaking directly to you” (p. 351). Additionally, Mayer asserted that using a foreign accent in narration might affect cognitive processing although he acknowledged that some learners may prefer a machine-generated voice or one with a foreign accent. He also suggested that learners could adapt to a machine voice or foreign accent with continued exposure.

Another principle based on social cues is Mayer’s (2014b) *embodiment principle*. This principle states, “people learn more deeply when on-screen agents display humanlike gesturing, movement, eye contact, and facial expressions” (p. 345). Based on the existing research, Mayer concluded that there is moderate evidence for using higher, more human-like levels of embodiment with on-screen, or pedagogical, agents; however, the author acknowledged that the effects are lessened when high-level embodiment is accompanied by other negative social cues such as a machine-generated voice.

One final principle based on social cues is Mayer’s (2014b) *image principle*. According to the image principle, including an image of the narrator in a multimedia presentation does not facilitate deeper learning; however, this principle is still being investigated because, as Mayer indicated, the conclusion was based on research studies that used low embodiment on-screen images such as, “a static image, a talking head, or a cartoon character that does not engage in much humanlike gesturing, movement, eye contact, or facial expression” (p. 360).

The *segmenting principle* refers to Clark and Mayer’s (2011) recommendation to break complex instruction into smaller, more manageable parts. The authors suggested manageable segments are “parts that convey just one or two or three steps in the process or procedure or describe just one or two or three major relations among the elements” (p. 209).

Clark and Mayer’s (2011) *pretraining principle* involves teaching key concepts and facts prior to presenting multimedia instruction. By doing so, learners have more working memory available to process the instructional material rather than being overloaded by unfamiliar terms or facts.

Although it does not always have its own chapter in some textbooks, *signaling* has been and remains a main principle of multimedia instructional design (R. Mayer, personal communication, January 26, 2019). Van Gog (2014) suggested that the signaling principle refers to improved learning from multimedia instruction when “cues are added that guide learners’ attention to the relevant elements of the material or highlight the organization of the material” (p. 263). According to Clark and Mayer (2011), “Signaling includes using headings, bold, italics, underlining, capital letters, larger font, color, white space, arrows, and related techniques to draw the learner’s attention to specific parts of the display or page” (p. 173). Mayer and Fiorella (2014) asserted that signaling reduces extraneous cognitive load because it draws the learner’s attention to the relevant material, which reduces the need to process extraneous material.

Furthermore, van Gog suggested, “signaling not only may prevent extraneous load by preventing processing of less relevant information, but may simultaneously foster germane load, by facilitating the organization or integration of essential material” (p. 264).

Cognitive-affective theory of learning with multimedia. Building on CTML, Moreno and Mayer (2007) presented a cognitive-affective theory of learning with media (CATLM), which was designed to include newer media, interactivity, and instructional approaches such as virtual reality and case-based learning. CATLM is based on the previously mentioned cognitive research assumptions, but, as shown in Figure 2, it also incorporates aspects of motivation and affect as well as metacognition.

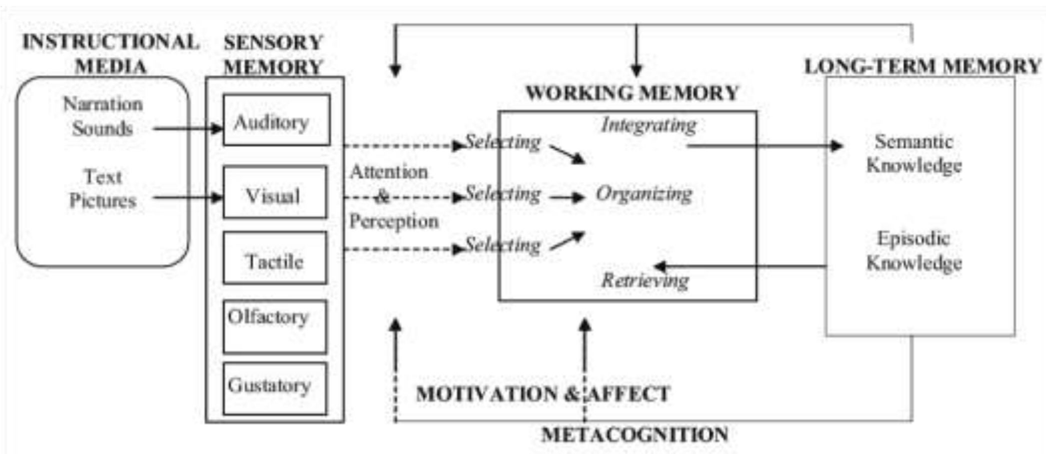


Figure 2. A cognitive-affective model of learning with media. From “Interactive Multimodal Learning Environments” by R. Moreno, and R. E. Mayer, 2007, *Educational Psychology Review*, *19*(3), p. 314. Copyright 2007 by *Educational Psychology Review.* Reprinted with permission (see Appendix A).

According to Moreno and Mayer (2007), motivational factors mediate learning by increasing or decreasing cognitive engagement. The authors stated, “When learners lack motivation they may fail to engage in generative processing even when cognitive capacity is available” (p. 315). Furthermore, Moreno and Mayer asserted, “learners may also use their metacognitive skills to regulate their motivation and cognitive processing during learning" (p. 314). Bruning, Schraw, and Ronning (1999) stated, "Educational psychologists became interested in metacognition because it was apparent that good learners know a lot about their own thinking and memory and use this information to regulate their learning" (p. 102). Mayer (2014a) echoed this sentiment stating, "in addition to being able to engage in appropriate cognitive processing during multimedia learning, successful learners must want to engage in appropriate cognitive processing (i.e., motivation) and know how to manage their cognitive processing (i.e., metacognition)" (p. 65).

Moreno and Mayer's (2007) CATLM is in alignment with the CTML, but also takes affective factors into account. As such, Moreno and Mayer outlined additional multimedia design principles and expanded on previous CTML principles with attention to their affective implications. The additional CATLM principles are discussed in more detail below.

According to Moreno and Mayer's (2007) *guided activity principle*, "Students learn better when allowed to interact with a pedagogical agent who helps guide their cognitive processing" (p. 316). According to the authors, guided activity occurs when learners interact with a pedagogical agent that facilitates essential and generative processing by encouraging learners to actively select, organize, and integrate new information. The authors asserted this processing leads to deeper understanding than when learners participate in pure discovery learning, which could increase extraneous cognitive load. Kirschner, Sweller, and Clark (2006) concluded that, "After a half-century of advocacy associated with instruction using minimal guidance, it appears that there is no body of research supporting the technique" (p. 83). Clark and Mayer (2011) referred to guided learning as part of whole-task instruction and suggested, “Learners are supported during the problem-solving episode to avoid mental overload” (p. 350).

According to the *reflection principle*, students learn better when they reflect on correct answers during meaning making. In a series of studies, Moreno and Mayer (2007) had learners reflect by explaining correct answers. The authors found that, “Adding reflection to an interactive environment does not significantly improve their learning, presumably because interactivity already primes the cognitive processes of organizing and integrating” (p. 127); however, “students who learn from noninteractive multimedia significantly increase their retention and far transfer with reflection techniques” (p. 127). Consequently, Moreno and Mayer concluded that adding this kind of reflection is particularly advantageous with passive instruction or if interactivity is superficial.

The *feedback principle* refers to the dependence of learning on students’ prior knowledge and quality of feedback. For example, Moreno and Mayer (2007) stated, “novice students learn better with explanatory rather than corrective feedback alone” (p. 318). The authors explained that corrective feedback relays to students only whether their answers are correct or incorrect, whereas explanatory feedback includes an explanation regarding why the answers are correct or incorrect. Moreno and Mayer asserted that explanatory feedback reduces extraneous processing when a mental model is not provided to students.

The *pacing principle* refers to the effect of learner control of the presentation pace and/or sequence on learning outcomes. For example, if an animation is presented too quickly, learners may not have enough time to process the material, which will affect learning outcomes; therefore, allowing learners to control the pace of complex multimedia instruction can provide them the necessary time to process the material. Additionally, Mayer, Dow, and Mayer (2003) found that being able to control the pace and order of instruction "reduces cognitive load by allowing learners to digest and integrate one segment of the explanation before moving on to the next" (p. 810). Moreno and Mayer (2007) asserted that allowing novice students to control the pace of instruction could reduce the amount of information learners must process in working memory at one time. The authors suggested dividing content into smaller segments that allow learners, "to select words and images from one segment and organize them before moving to the next segment" (p. 320) as another possibility to provide learners adequate time for deeper processing.

The *pretraining principle* refers to improved learning outcomes when necessary prior knowledge is activated or learners are provided with relevant prior knowledge before the instruction is presented. Moreno and Mayer's (2007) pretraining principle appears to be the same principle outlined by Clark and Mayer (2011); however, Moreno and Mayer discussed pretraining in the context of non-interactive versus interactive multimedia instruction. The authors asserted that pretraining is a type of interactivity, "where students may ask to learn about the component of a system or a pedagogical agent may offer an explanation about the system components, if needed" (p. 320-321). Moreno and Mayer suggested, "pretraining helps students engage in generative processing by showing them which pieces of prior knowledge they should integrate with incoming information" (p. 320).

Summary of the theoretical framework. As Clark and Mayer (2011) asserted, "e-learning courses should be constructed in light of how the mind learns and experimental evidence concerning e-learning features that best promote learning" (p. 29). Mayer (2014a) asserted the knowledge in long-term memory is analogous to what Sweller (1994) referred to as schemas and expertise results from schema acquisition and automation (Van Merriënboer & Kester, 2014). According to Sweller (1999), schemas "provide the means of storing huge amounts of information in long-term memory (p. 11). Mayer (2014a) stated, "As new knowledge is constructed in working memory, it may be stored in long-term memory as prior knowledge to be used in supporting new learning" (p. 59). The constraints of working memory led Sweller (1994), to advance the cognitive load theory, which "suggests that instructional techniques that require students to engage in activities that are not directed at schema acquisition and automation, frequently assume a processing capacity greater than our limits and so are likely to be defective" (p. 299). By fostering generative processing, which is required to make sense of the essential material, instructional designers also foster schema acquisition and automation, or germane cognitive load.

From cognitive load theory, Mayer and Moreno (2002) introduced the CTML, and later the CATLM, which rest on the assumptions that people have separate verbal and visual information processing channels, these channels have limited processing capacity, and meaningful learning occurs when people actively select and organize relevant information and integrate it with existing knowledge. Building on those assumptions, Clark and Mayer (2011) advanced principles for managing processing demands of learning: multimedia, contiguity, modality, redundancy, coherence, personalization, segmenting, and pretraining principles. Additionally, Moreno and Mayer (2007) outlined the guided activity, reflection, feedback, pacing, and pretraining principles based on the assumptions that learning is mediated by motivational factors that increase or decrease cognitive engagement and metacognitive factors that regulate cognitive processing and affect. Furthermore, the amount of material learned from certain media may be affected by learners' prior knowledge and abilities. According to Mayer (2014c), the coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles are designed to reduce extraneous processing while the segmenting, pretraining, and modality principles are designed to manage essential processing and the personalization, voice embodiment, and image principles are designed to foster generative processing; therefore supporting cognitive processes to maximize learning outcomes.

Situated Learning Theory

Brown, Collins, and Duguid (1989) discussed situating learning into the context of where it will be used. These authors asserted that conceptual knowledge and its implementation in practical situations are disconnected in traditional instruction. Therefore, learning requires authentic activity rather than contrived activities, which are normally presented in schools. Brown et al. stated, "People who use tools [knowledge] actively rather than just acquire them . . . build an increasingly rich implicit understanding" (p. 33). According to the authors, in traditional education settings, "students may pass exams . . . but still not be able to use a domain's conceptual tools in authentic practice" (p. 34). To address the disconnect between knowing and doing, the authors suggested the use of cognitive apprenticeships to contextualize learning using the social and physical context of the instruction. Using cognitive apprenticeship, students should progress from specific authentic activities to more general principles by generating their own paths to solutions. The authors stated, “learning methods that are embedded in authentic situations are not merely useful; they are essential” (p. 37).

Choi and Hannafin (1995) also recognized the tendency of formal instruction to decontextualize knowledge. The authors stated, “Typically, formal education emphasizes general, abstract decontextualized knowledge which is often difficult to transfer to real-life situations” (p. 57). Instead of such artificial activities they stressed an apprenticeship concept, which “emphasizes the importance of experiential activity in learning, and highlights the inherently context-dependent, situated, and enculturated nature of learning” (p. 60). The authors asserted, “Situated cognition emphasizes the importance of context in establishing meaningful linkages with learner experience and in promoting connections among knowledge, skill, and experience” (p. 54). Choi and Hannafin suggested situated learning anchors knowledge and skills in realistic contexts. These realistic contexts include authentic tasks, which “are more likely to become self-referenced and purposefully engaged by learners” (p. 56). The authors stated, “Context provides the framework for learning, but content determines its authenticity and veracity” (p. 57). Further, Choi and Hannafin indicated, “situated learning environments emphasize the use of diverse concrete instances in authentic contexts” (p. 59) to create knowledge and skill that are both specific and general.

Herrington and Oliver (1999) qualitatively examined pre-service teachers’ higher-order thinking while using an interactive multimedia intervention based on situated learning theory. The authors video-recorded four groups with two students using the intervention and analyzed their discussions for specific levels to indicate higher- and lower-order thinking. The authors found that the majority of discussion fell in the higher-order talk; however, "there were many differences in the extent and levels of the various forms of higher-order thinking evident in the talk" (p. 6). The authors suggested the instrument might not have been precise enough to accurately measure student cognition.

Lee and Butler (2003) used CD-ROMs to deliver instructional resources to situate sixth-grade students as meteorologists, which allowed them to collect data, compare weather data from different regions, interpret weather maps/images, and forecast. Based on the results, the authors suggested, "not all authentic situations are appropriate for the development of students' understandings of scientific knowledge" (p. 27). Lee and Butler asserted real world situations "must map closely to students' content understandings and curricular activities" (p. 27).

Oliver and Herrington (2000) outlined nine critical elements for designing learning environments based on situated cognition and learning, which are authentic contexts, authentic activities, access to expert performances and modeling of processes, multiple roles and perspectives, collaborative construction of knowledge, reflection to enable abstraction, articulation to enable making tacit (implicit) knowledge explicit, teacher coaching and scaffolding at critical times, and authentic assessment of learning within the tasks. The authors suggested, "Situated learning can be achieved to some degree by the inclusion of any of these elements in a learning setting" (p. 183), but the challenge is to explore how all nine elements could be incorporated to act together to support learning. Oliver and Herrington discussed their application of these principles using problem-based learning (PBL) and reported increased student motivation, critical thinking skills and deeper understanding.

More recently, Sung, Hwang, and Yen (2015) compared the effect of a contextual game-based instructional intervention against a conventional e-book on learning motivation, learning performance, and problem-solving abilities. Results showed the intervention improved learning motivation, learning achievement, and problem-solving abilities. The authors found a significant two-way interaction and suggested the game-based intervention benefited students with higher motivation in terms of advanced knowledge.

Clark and Mayer (2011) stated, "From the plethora of media comparison research conducted over the past sixty years, we have learned that it's not the delivery medium, but rather the instructional methods that cause learning" (p. 14). Clark and Mayer asserted that although there are pitfalls, e-learning offers the promise of customized training, learning engagement, multimedia, and accelerated expertise through scenarios.

Scenario-based Learning

Scenario-based learning (SBL) grew out of situated learning theory (Errington, 2011), which posits that meaning comes from the context of use (Brown et al., 1989). Carroll (1999) defined scenarios as stories about people and their activities. Carroll suggested that natural learning does not occur using the systems approach of breaking instruction down into hierarchical tasks and subtasks stating, "People can learn in spite of such an approach, but only because they are so adaptable" (p. 7). Carroll asserted that, "People want to learn in a realistic context; they want to be able to use what they already know to critically evaluate new knowledge and skills they encounter" (p. 7). Clark

(2016) asserted that SBeL is an instructional method for accelerating the growth of expertise. Clark discussed the lack of a common taxonomy of terms in the instructional design field. For example, Clark stated, “SBeL can be called problem-based learning, case-based learning, immersive learning, or whole task learning, just to name a few” (p. 51).

Clark and Mayer (2011) also suggested that scenario-based instruction is synonymous with whole-task, case-based, problem-based, and immersive instruction. According to these authors, “*whole task instruction* begins the lesson with an authentic work assignment and integrates the needed knowledge and skills in the context of working on that assignment” (p. 345, emphasis in original). This approach contrasts with part-task (also called directive) instruction, “in which content is broken into small segments, prerequisite knowledge is usually taught first, and frequent practice with feedback helps learners build skills gradually” (p. 345).

Lim, Reiser, and Olina (2009) examined undergraduate pre-service teachers’ (N = 51) acquisition and transfer of Excel software skills using whole-task (SBL) and part-task instruction. The authors also examined the possible interaction between learners’ prior knowledge and the instructional approaches. Prior to the instruction, participants completed a pretest to determine whether participants could perform specific Excel skills and were placed in higher or lower prior knowledge groups accordingly. Both instructional methods included two 60-minute lessons separated by five minutes and participants in both groups had access to the instructor throughout the treatment conditions. In the first segment of each lesson under both treatment conditions, the instructor provided an overview of the topics, the lesson, concepts and skills, and examples of Excel grade books that were designed using those skills. Following the first segment, the instructor presented and demonstrated constituent skills of preparing an Excel grade book to participants in the part-task group. These participants were taught 15 “fairly basic skills” (p. 66) during the first lesson and seven more advanced skills during the second lesson. Following this demonstration, participants practiced using the skills. At the end of the second lesson, participants in the part-task group prepared an Excel grade book.

In contrast, participants in the whole-task group were taught the constituent skills in the context of the whole task rather than as isolated procedures. During both whole-task lessons, participants in the Lim et al. (2009) study created Excel grade books as the instructor demonstrated how to create them and then created un-weighted and weighted grade books in the first and second lesson respectively. Participants in both groups completed a part-task achievement test, whole-task achievement test, and a transfer test. Differences among participants on these three dependent variables were examined using a two-way (instructional method × prior knowledge) multivariate analysis of variance (MANOVA).

To complete the part-task achievement test, participants in both groups performed 16 separate part-tasks. One portion of the test was objective while two raters graded the other portion using a rubric and the maximum score was 34. According to Lim et al. (2009), participants in both groups performed well on the part-task achievement test with mean scores of 91% ($M = 30.9$, $SD = 3.14$) and 92% ($M = 31.3$, $SD = 1.95$) for participants in the whole-task and part-task groups respectively. Results of the MANOVA showed a significant overall main effect [$F(3, 45) = 6.38$, $p < 0.05$, $\eta^2 = 0.3$]

for instructional approach (part- or whole-task) on skill acquisition and transfer. However, a post-hoc ANOVA with a Bonferroni adjusted alpha level of 0.017 found no significant main effect of instructional approach (part- or whole-task) on skill acquisition as measured by the part-task achievement test. The authors found no significant main effect of prior knowledge on participants' performance on the part-task achievement test. Additionally, the authors did not observe a significant interaction between prior knowledge and instructional approach. The results indicate that both part-task and whole-task instruction produced equivalent results across differing prior knowledge levels when participants were assessed on their ability to perform part-task skills.

The whole-task achievement test required participants to create an Excel grade book that incorporated specific features, which had been covered in both treatment conditions. Similar to the part-task achievement test, a portion of the whole-task achievement test was graded by two raters using a rubric and had a maximum score of 36. As with the part-task achievement test results, Lim et al. (2009) noted participants in both groups performed well on the whole-task achievement test with mean scores of 89% (M = 32.2, SD = 5.0) and 80% (M = 28.7, SD = 5.72) for participants in the whole-task and part-task groups respectively. Because the initial MANOVA detected a significant main effect of instructional approach on participants' skill acquisition (part-task and whole-task achievement tests), the authors conducted a follow-up ANOVA using a Bonferroni adjusted alpha level of 0.017, which found participants in the whole-task group scored statistically significantly higher [$F(1, 47) = 6.12$, $p < 0.017$, $\eta^2 = 0.12$] than participants in the part-task group on the whole-task test. Lim et al. also reported a moderate effect size of d = 0.71 (Myers, Well, & Lorch, 2010). These results indicate that participants who were taught using the whole task approach were also better able to complete the whole task than those who were taught only partial tasks. Prior knowledge was not found to have a main effect or a significant interaction effect on the whole-task achievement test; however, the authors suggested that the pretest may not have accurately identified higher and lower prior knowledge and that the whole-task achievement test of creating an Excel grade book might have been simple enough for the participants to minimize the possible interaction between prior knowledge and achievement.

The transfer test involved using the acquired skills to create a budget in Excel that included specific features from the skills that were taught during both treatment conditions. Participants in the whole-task group had a mean score of 86% (M = 30.9, SD = 4.36) and the participants in the part-task group had a mean score of 68% (M = 24.6, SD = 6.67) with the possible total score being 35.5. The results of the follow-up ANOVA showed that participants in the whole-task group performed significantly better [$F(1, 47) = 15.87$, $p < 0.017$, $\eta^2 = 0.25$] on the transfer test than students in the part-task group. A large effect size estimate (d = 1.14) was also reported (Myers et al., 2010). A follow-up ANOVA also revealed that participants with higher prior knowledge (M = 30.0, SD = 4.8) scored significantly higher [$F(1, 47) = 6.57$, $p < 0.017$, $\eta^2 = .12$] than those with lower prior knowledge (M = 25.9, SD = 7.04). However, no interaction was found between prior knowledge and treatment condition, which would support the idea that learners with higher prior knowledge were better able to transfer the component skills. Overall, based on their findings, Lim et al. (2009) concluded that whole-task instruction may promote skill acquisition and transfer; however, the authors recommended future research "to determine whether these promising findings hold true across a wide range of cognitive skills and learners" (p.75).

Loyens, Jones, Mikkers, and van Gog (2015) examined the ability of problem-based instruction (*n* = 24) to promote conceptual change as compared to traditional, lecture-based instruction (*n* = 23) and self-study instruction (*n* = 24). The instruction was designed to address misconceptions about Newtonian laws, specifically the impetus theory, or how objects fall. Participants in all three groups completed a pretest prior to the treatments which each lasted 60 minutes. Participants in the PBL group were presented with a three-part problem presented on paper. Each of the three parts consisted of separate discussions among a group of friends regarding how objects fall. First, participants discussed possible explanations and solutions for 20 minutes, which yielded unanswered questions that were used to develop questions, or "learning issues" (p. 38) on which to base their self-study. Participants in the PBL group were then provided 20 minutes to study a text on impetus theory followed by another 20 minutes to discuss their subsequent answers to the learning issues, which was guided by a tutor. Following the guided discussion, the text and any participant notes were collected and participants completed the immediate posttest.

After completing the pretest, participants in the Loyens et al. (2015) self-study group studied the text describing impetus theory, which also included examples and the three-part problem given to participants in the PBL group. Self-study participants were provided 60 minutes to study the text and take notes, after which, participant notes were collected and the participants completed the immediate posttest. Participants in the lecture-based treatment group received the same information from the impetus theory text, but an instructor presented the information. The lecture was limited to 60 minutes to keep instruction time equal among the groups. Following the lecture, participants turned in their notes and completed the immediate posttest.

Loyens et al. (2015) administered a single instrument as the pretest, the immediate posttest, and the delayed posttest. The test required participants to view short video animations showing direction of movement and then draw lines on paper illustrating the path they expected falling objects to follow. A repeated measures analysis of variance (ANOVA) was employed to measure the change in student scores from pretest to immediate and delayed posttests. The results indicated participants in all three groups gained knowledge, but participants in the PBL group performed better than participants in the lecture-based and self-study groups on both the immediate and delayed posttests. Although the results appeared to indicate that participant scores between the immediate and delayed posttests dropped more for participants in the lecture and self-study groups than participants in the PBL group, paired samples *t*-tests showed none of the decreases were statistically significant. The authors concluded, "the present findings suggest that PBL students experienced greater levels of conceptual change from pretest to immediate posttest" (p. 39), which the authors contribute to the elements of PBL instruction that can facilitate conceptual change such as activating prior knowledge, engaging in discussion, and critically evaluating the material.

Landrigan (2010) explained the development of scenario-based online modules to help international students with the Issues, Rule, Application, and Conclusion (IRAC) method of approaching corporate law. The first module covered the Issues component of IRAC while a second module covered the Rule, Application, and Conclusion components. Both modules presented scenarios centered around a student employed part-time at a law firm. To complete the modules, students were required to investigate statutes and cases for advising a client about an appropriate course of legal action. The author reported feedback that students were not using the first module because it was viewed as too long, which "reinforced the suggestion . . . that the modules be accompanied by tutor interaction" (p. 525). Although no quantitative data were included in the article, the author concluded, "The findings confirmed the modules are of benefit when accompanied by active teaching to ensure that essential skills in Corporations Law be made explicit to all students" (p. 526).

In addition to using SBeL for teaching diagnosis and repair, compliance, and interpersonal skills (Clark, 2013), SBL has also been used within teacher education (Hursen & Fasli, 2017; Meldrum, 2011; Sorin, 2013; Yetik, Akyuz, & Keser, 2012) and health professions (Cook et al., 2008). Hursen and Fasli (2017) compared the influence of SBL and reflective learning on prospective teachers' ($N = 62$) achievement and professional self-competence. Participants were assigned to groups based on even or odd school numbers and completed a pretest to assess their knowledge of the content to be taught during a 12-week teacher practice course and a self-competence evaluation. Participants in the SBL group participated in discussions, research, and solution generation with regard to scenarios. The participants also formed small groups and prepared reports, which were presented to the class. Participants in the reflective learning group reported what they learned and answered reflective questions in learning writings. The participants also reflected on concept maps and lesson plans they generated over the 12-week course. After the treatments, all participants completed the achievement test and self-competence evaluation. The results of the achievement posttest indicated participants in the SBL group scored significantly higher than participants in the reflective learning group. Based on the results, Hursen and Fasli asserted, “the scenario based learning approach is more efficient for the academic achievement levels of the prospective teachers” (p. 275).

Meldrum (2011) presented a case study of the use of SBL for preparing pre-service physical education teachers to address potential issues they may face in the future. Participants identified driving factors and wrote scenarios based on the question, “What is my future as a professional over the next three years?” (p. 135). Participants were also separated into small groups and asked to create scenarios addressing issues such as how to increase girls’ physical education participation and how faculty might handle issues of limited space or facilities. Based on participant survey responses, the author concluded that SBL could encourage “students to think about difficult school-based questions and to attempt to provide solutions to them” (p. 142).

Sorin (2013) discussed the use of SBL for preparing early childhood education students using an action research approach. The SBL narratives were based on actual experiences of early childhood professionals, which were presented to the participants along with supporting documentation such as anecdotes and children’s work samples. The participants assumed the role of early childhood teachers and crafted possible solutions to the problem-based scenarios. In some instances, participants also role-played different characters in the scenario to provide additional drama as well as responsibilities for group work such as collecting specific information regarding particular characters. Further, the researcher enlisted student actors to film short, supplemental video clips.

Over time, the researcher also modified student assessment by transitioning from written reports about how to address the scenarios to a more varied approach such as writing articles to help parents support their children's learning or an acceptance speech for an inclusive practice award. Sorin concluded that using this SBL approach, "students seem to be gaining a better understanding of the issues explored in the scenarios and how they might react if they encounter a similar situation" (p. 80).

Yetik et al. (2012) examined the use of scenarios with (N = 37) preservice teachers using a pretest/posttest experimental design. Learners were randomly assigned to one of five schools to observe and then lecture as part of a Teaching Practice course. A Turkish version of a problem-solving inventory (PSI) was used as the pretest and posttest to assess students' self-perceived problem-solving attitudes and behaviors using a paired samples t-test. The learners were presented with one scenario a week for eight weeks, which were also discussed in face-to-face sessions, to help learners with potential problems they may encounter in the classroom. Yetik et al. concluded that SBL positively affected learners' perceptions of their problem-solving skills. The authors also asserted that presenting the course topics using SBL increased preservice teachers' problem-solving skills; however, no learning outcomes were reported although Yetik et al. stated, "generated solutions according to presented problems were assessed with students at the last week of the study" (p. 164).

Yarnall, Toyama, Gong, Ayers, and Ostrander (2007) reported on a three-year case study of scenario-based technical instruction at five community colleges. According to the authors, "The objectives of the project were to develop new course materials, disseminate them, test their success for student learning, study their feasibility of implementation, and examine their effectiveness for moving a greater number of students into professional ranks" (p. 587). During the first year, six scenario-based tasks for a Practicum in Enterprise Security course were pilot-tested over a 12-week quarter followed by revisions based on student and instructor feedback. The program evaluation group also measured students' technical content learning outcomes and professional skills; however, the authors focused instead on student and instructor feedback. Three themes emerged from the survey, interview and focus group data: quality of curriculum materials, quality of assessment, and difficulty of implementation. The authors found that, "community college instructors and students wanted more structure to clarify each scenario-based task's learning goals and help with the challenges of teamwork" (p. 598), but the scenario-based curriculum was well received by both students and instructors after increasing the structure. The authors also asserted, "Our work also indicates that scenario-based curriculum is more accessible to students with past work experience and somewhat intimidating to technical novices" (p. 599).

Health professions. Extensive research has been conducted on Internet-based instruction in the health professions as evidenced by the 201 studies analyzed by Cook et al. (2008). The authors examined studies comparing Internet-based instruction to no instruction and Internet-based instruction to traditional instruction. Because the meta-analysis was limited to Internet-based instruction studies, other forms of computer-based instruction, such as CD-ROM tutorials, were not included. The results of this meta-analysis showed inconsistent results for Internet-based instruction. The results of some studies favored Internet instruction while others favored traditional instruction. The authors asserted, "Internet-based instruction appears to have a large effect compared with no intervention and appears to have an effectiveness similar to traditional methods” (p. 1195). The authors concluded, “Internet-based instruction is associated with favorable outcomes across a wide variety of learners, learning contexts, clinical topics, and learning outcomes” (p. 1195); however, the authors cautioned that the inconsistencies among studies as well as publication bias could have affected the results. Because of the inconsistent results, the authors suggested that rather than further comparison research, the focus should now be on how best to implement and under what conditions Internet-based instruction works best. Choules (2007) also suggested, "Comparison with more ‘traditional’ teaching is unhelpful. The future is to look at what elearning can do and use it to its strengths" (p. 216).

Within the health professions field, SBL has been used as a substitute for more lab time due to time constraints, cost, and safety hazards in the physical laboratory setting (Breakey, Levin, Miller, & Hentges, 2008; Lehmann, Bosse, Simon, Nikendei, & Huwendiek, 2013). Scenario-based e-learning has also been used to teach soft skills such as patient management (Carpenter et al., 2003; Choules, 2007; Clark, 2013; Lehmann et al., 2013).

Breakey et al. (2008) created an SBeL genetics activity, they termed PBL, using interactive software. The scenario-based module was based on a fictitious “chocolate monster” to illustrate mutations and isolate new mutant animals by conducting breeding experiments. The authors did not provide quantitative results, but reported that, “Sixty percent of students (16 of 27) agreed or strongly agreed that the online PBL was a useful addition to the course” (p. 1155). In addition, 78% agreed or strongly agreed that corrective feedback was useful. The authors also stated that students enjoyed the interactive aspect of it.

Carpenter et al. (2003) compared motivational interviewing learning differences between participants who completed an online role-play simulation versus those who were instructed to spend at least 45 minutes on the national smoking cessation guidelines website while focusing on patients who are unwilling to quit ($N = 28$). The authors found higher scores on the smoking cessation skills and techniques posttest for participants who completed the simulation. The authors concluded that the results “support the idea that healthcare professionals can learn clinical skills from computer-based training software” (p. 158).

Lehmann et al. (2013) presented their work using virtual patients (VPs) to prepare students for pediatric clinical skills laboratories. They explained that skills laboratories require several tutors and time slots to work with students in small groups and the laboratory time is usually spent providing instruction and demonstrations prior to repeated practice and feedback. The VPs were used to provide students instruction and demonstration prior to laboratory time. By preparing students ahead of time, laboratory time could be used more efficiently for repeated practice. According to the authors, students and tutors responded well to the multimedia VPs stating, “Students felt well prepared by VPs and thus could efficiently use the skills laboratory time for practical training” (p. 6).

Summary of scenario-based learning. From the literature, it is apparent that SBeL is not a universal answer to all learning situations and outcomes (Clark, 2016); however, Lim et al. (2009) found that learners who were taught using a whole-task approach were better able to complete a whole-task assessment than those who were taught in parts, which could indicate that SBeL instruction promotes learning and transfer. Further, Landrigan (2010) found that SBeL worked better with the addition of active teaching techniques. More recently, Loyens et al. (2015) suggested learners in a PBL setting showed higher knowledge gains between a pretest and immediate posttest than those taught using traditional, lecture-based or self-study instruction, and Hursen and Fasli (2017) concluded that SBL was more efficient for academic achievement of prospective teachers than reflective learning. Additionally, there are also varying implementations of SBL and SBeL. For example, Hursen and Fasli used scenarios with group discussions, research, and reports while others, such as Landrigan, used stand-alone modules; however, according to Clark (2013) a key factor for SBeL success "is sufficient guidance to minimize the flounder factor" (p. 6). Citing the results of studies by Lim et al. (2009) and Loyens et al. (2015), Clark (2016) stated, "scenario-based learning can lead to better transfer to tasks different from those used in the training and to greater long-term retention" (p. 55); however, Clark (2016) also cautioned, "we need more research before extending these conclusions to all situations" (p. 55).

Feedback

According to Hattie and Gan (2011), feedback is information about an individual's performance or understanding provided to lessen the gap between the individual's current performance or understanding and the desired performance or understanding. These authors reviewed 12 meta-analyses examining classroom feedback and found feedback to be in the top ten factors influencing achievement; however, "The variance of effects . . . was considerable, indicating that some types of feedback are more powerful than others" (p. 249). Azevedo and Bernard (1995) extracted effect sizes from 14 studies examining feedback. According to the authors, "Feedback has to be regarded as one of the most critical components of computer-based instruction, its objective being to provide students with appropriate responses thus allowing them to rectify learning difficulties" (p. 120). Lechermeier and Fassnacht (2018) stated, "The effects of performance feedback on individuals' behavior traditionally have attracted the attention of scholars of a variety of disciplines" (p. 145). The field of feedback is large with hundreds of published research studies exploring the topic of feedback as well as its connection to learning and performance (Shute, 2008). Although there is a multitude of research on feedback, the findings are varied and inconsistent (Shute, 2008); however, the extensive research on feedback has produced compelling evidence for its effect on learning outcomes (Hattie & Gan, 2011). Van der Kleij, Feskens, and Eggen (2015) asserted, "Many researchers have noted that the literature on the effects of feedback on learning provides conflicting results" (p. 476). Feedback can be classified by type, level, and timing (Van der Kleij, Timmers, & Eggen, 2011). For the purposes of this study, the review of the literature will organize feedback by types, timing, and delivery methods, which will be addressed next.

Feedback types. Jaehnig and Miller (2007) outlined that feedback messages can be delivered as knowledge of results (KR), knowledge of correct response (KCR), and elaboration feedback (EF). According to Jaehnig and Miller, KR feedback indicates whether an answer is correct or incorrect, KCR includes the correct answer and often whether the learners' answer was correct, EF includes the correctness of the response, the correct answer, and additional information or an explanation. The KR feedback is analogous to Johnson and Priest's (2014) definition of corrective feedback and Shute's (2008) verification feedback in which learners are shown only whether their answers are correct or incorrect. Similarly, KCR refers to confirming the correctness of responses, but also providing the correct answer (Van der Kleij et al., 2015). Jaehnig and Miller suggested EF is also referred to as extended feedback. Shute explained that EF "can (a) address the topic, (b) address the response, (c) discuss the particular error(s), (d) provide worked examples, or (e) give gentle guidance" (p. 158). For example, according to Shute, EF may include an explanation of why an answer was wrong in addition to indicating the correct answer.

Feedback types can go by different names with similar definitions. For example, Clark (2013) suggested instructional feedback is the same as corrective feedback. According to Clark, instructional or corrective feedback consists of providing learners with "the accuracy of a solution, along with an explanation" (p. 107); however, this type of feedback fits more with Jaehnig and Miller's (2007) elaboration feedback and Johnson and Priest's (2014) explanatory feedback in which the learners are provided "with a principle-based explanation of why his or her answer was correct or incorrect" (p. 450).

Johnson and Priest (2014) asserted that explanatory feedback has been found to be more effective than corrective feedback for novice learners. The authors asserted, "providing corrective feedback only is a minimally guided method of instruction that merely informs the learner that he or she is right or wrong, with no additional information" (p. 451). The authors explained that, because no additional information is provided, learners must search for additional information and attempt to correct their misunderstandings, which increases extraneous cognitive load.

Moreno (2004) compared the efficacy of explanatory feedback and corrective feedback delivered verbally by an on-screen pedagogical agent in a botany multimedia lesson. The author conducted two experiments designed to test a *guided feedback hypothesis*, “according to which, discovery-based environments that guide students’ discovery by means of explanatory feedback promote deeper learning than those that present corrective feedback alone” (p. 100). In both the explanatory and corrective feedback treatments, the pedagogical agent (a) asked learners to select a suitable plant for a specific environmental condition (rain, wind, etc.), (b) provided information about the correctness of the response, and (c) provided learners with the right answer before moving to the next step; however, learners in the explanatory condition were also provided explanations regarding why answers were correct or incorrect prior to receiving the correct answer. Moreno found the EF treatment to be statistically significantly more effective for retention and transfer, though the results were admittedly limited to a botany lesson, learners who were unfamiliar with botany, and with scientific explanation feedback. The author concluded, “unguided instructional methods impose heavy performance demands on novice learners thus leaving little room for germane load” (p. 110).

Jaehnig and Miller (2007) asserted that, “The term feedback has also been used to describe potentially punishing consequences for incorrect responses or added incentives for correct responses” (p. 222). Discussing rewards as incentives for learning, Hattie and Gan (2011) stated, “Feedback as extrinsic rewards often led students to place more emphasis on incentives, which result in greater surveillance, evaluation and competition, rather than enhanced engagement in learning” (p. 252).

According to Jones (2018), "Pointing out errors or deficiencies without providing information that could prevent future repetitive errors does little to promote improvement" (para. 2). According to Hattie and Timperley (2007), "[Feedback] is most powerful when it addresses faulty interpretations, not a total lack of understanding" (p. 82). The authors also stated, "Those studies showing the highest effect sizes involved students receiving information feedback about a task and how to do it more effectively. Lower effect sizes were related to praise, rewards, and punishment" (p. 84). Similarly, Clark and Mayer (2011) asserted that incorporating an explanation with feedback "provides a much better opportunity for learning" (p. 263) because "A missed question is a teachable moment" (p. 263).

On a more general level, feedback can also be delivered as formative or summative assessment (Shute, 2008). Shute (2008) suggested summative feedback is delivered after completion of a quiz or test while formative feedback is "information communicated to the learner that is intended to modify his or her thinking or behavior for the purpose of improving learning" (p. 154). According to Xu, Kauer, and Tupy (2016), formative feedback is "aimed at modifying thinking or behavior to improve learning" and should judge the response rather than the student (p. 151). Stufflebeam and Shinkfield (2007) suggested formative evaluation provides feedback for improvement while summative evaluation is intended to provide an overall judgment. According to Black and Wiliam (2010), "The task of assessing pupils summatively for external purposes is clearly different from the task of assessing ongoing work to monitor and improve progress" (p. 89). These authors also suggested, "There is a body of firm evidence that formative assessment is an essential component of classroom work and that its development can raise standards of achievement" (p. 90). Furthermore, Shute (2008) suggested, "The main goal of formative feedback–whether delivered by a teacher or computer, in the classroom or elsewhere–is to enhance learning, performance, or both, engendering the formation of accurate, targeted conceptualizations and skills" (p. 175). Consequently, the focus of the current study is on the use of formative feedback to determine whether its use affects learning outcomes.

Feedback delivery. Feedback can be verbal, written, or computer-based (Azevedo & Bernard, 1995; Hattie & Timperley, 2007; Jaehnig & Miller, 2007). Feedback can be delivered informally or formally and can be delivered by the instructor, peers, or self (Hattie & Timperley, 2007). Hattie and Timperley (2007) stated:

> A teacher or parent can provide corrective information, a peer can provide an alternative strategy, a book can provide information to clarify ideas, a parent can provide encouragement, and a learner can look up the answer to evaluate the correctness of a response. (p. 81)

Feedback timing. Aside from different feedback types, feedback can also be delivered at different times (Jaehnig & Miller, 2007). According to Clariana, Wagner, and Murphy (2000), immediate and delayed feedback differ in both timing as well as stimulus exposures. For example, with immediate feedback, learners receive the feedback near the same time they respond to the question/prompt; whereas, "Typically, delayed feedback clearly presents at least two exposures with each item, the first during the learner's initial response, and the second after some time delay when the feedback containing the stimulus and correct response are presented" (p. 6). The amount of delay can also vary (Shute, 2008). According to Shute (2008), immediate feedback can refer to directly after responding to a single item or an entire quiz; whereas, delayed feedback "is usually defined relative to immediate [feedback], and such feedback may occur minutes, hours, weeks, or longer after the completion of some task or test" (p. 163).

Summary of feedback literature. Although there are hundreds of studies researching feedback and its effects on learning, the results of these studies have been inconsistent (Shute, 2008). Lechermeier and Fassnacht (2018) echoed this sentiment stating, "Although an extensive body of research has stressed its importance, a conclusive overall picture on feedback characteristics effects is missing" (p. 145). Despite the inconsistent and inconclusive results, Shute (2008) suggested, "providing feedback that is specific and clear, for conceptual and procedural learning tasks, is a reasonable, general guideline" (p. 158).

Burkšaitienė (2011) suggested, "delivering feedback, especially critical feedback, may cause strong emotional reactions (including defensiveness, rejection of feedback, perception of feedback as a personal attack, threatening one's ego) all of which may hinder learning" (p. 40). Consequently, Burkšaitienė suggested that motivating and engaging feedback should be provided regularly both in oral and written forms in an atmosphere that fosters learning. The author recommended the feedback should also be formative, constructive, clear and easy to understand as well as encourage reflection. In addition to its effects on learning outcomes, feedback has also been linked to learner confidence (Clark, 2013; Hattie & Gan, 2011; Shute, 2008), which is one aspect of learner motivation (Keller, 2010). Azevedo and Bernard (1995) stated, "The motivating aspects of feedback are often encountered in the literature" (p. 112) while Nicol and Macfarlane-Dick (2006) suggested, "Research shows that feedback both regulates and is regulated by motivational beliefs" (p. 201). Keller (2010) asserted, "Feedback can be a powerful tool for building positive motivation and also for killing people's motivation" (p. 165).

Motivation and the ARCS Model

Providing learners with corrective feedback as well as how to remedy any mistakes can help build learner confidence (Keller, 2010). According to Keller (2010), confidence is a psychological characteristic that influences motivation, learning, performance, and attitudes. Attention (curiosity) and relevance (motives) are two other psychological characteristics that Keller suggested influence motivation, learning, performance, and attitudes. Additionally, according to Keller, "in order for [learners] to have a continuing desire to learn, they must have feelings of satisfaction with the process or results of the learning experience" (p. 46); therefore, satisfaction comprises the fourth category of Keller's ARCS model.

Keller (2010) asserted that attention relates to not only capturing the learner's attention, but also sustaining it. Additionally, Keller suggested learner attention is related to both motivation and learning. Keller stated, "In motivation, the issue is with how to **stimulate** and **sustain** the learner's attention . . . . In learning, the concern is with how to **direct** the learner's attention to the concepts, rules, skills, or facts to be learned" (p. 76, emphasis in original).

Relevance refers to how useful learning is perceived (Keller, 2010). The more useful a goal is perceived by the learner, the more likely an individual is to pursue its accomplishment presuming the individual also perceives the goal to be attainable. According to Keller (2010), "Before students can be motivated to learn, they will have to believe that the instruction is related to important personal goals or motives and feel connected to the setting” (p. 45).

Keller (2010) stated that confidence generally refers to “people’s expectancies for success in the various parts of their lives” (p. 135). The author suggested that a goal of motivational design with respect to confidence is to build expectancy for success. To do this, Keller recommended, “helping students understand what is expected of them and how to maximize their likelihood of success” (p. 164).

Satisfaction refers to individuals’ feelings and attitudes toward their performance and their expectations (Keller, 2010). Designing for motivational learning includes facilitating feelings of satisfaction and accomplishment (Keller, 2010). According to Keller (2010), such design could be supported by “using exercises that are authentic with an optimal challenge level, by providing feedback on results, and by grading in a manner that is fair and consistent with a stated set of criteria” (p. 188).

Keller (2010) suggested the ARCS model illustrates how a learner’s overall performance is affected by the amount of effort put into accomplishing a goal, which is influenced by a combination of motivation and the learner’s knowledge and skills as shown in Figure 3.

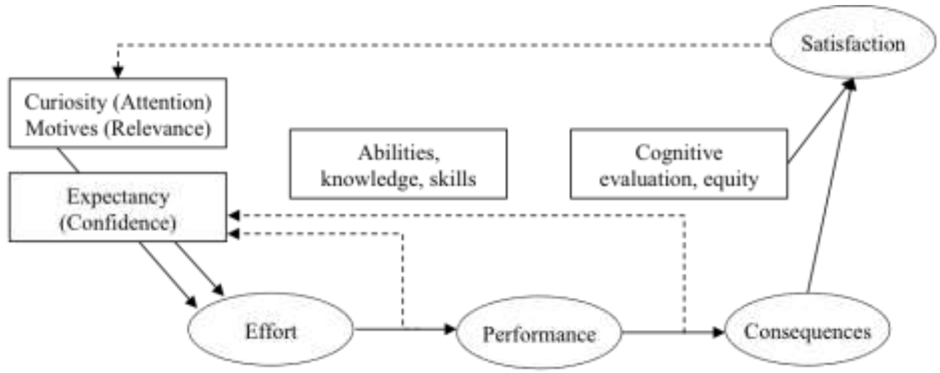


Figure 3. A model of motivation and performance. Adapted from *Motivational Design for Learning* (1st Ed., p. 6), by J. M. Keller, 2010, New York, NY: Springer. Copyright 2010 by Springer. Adapted with permission (see Appendix A).

According to Keller (2010), although a tremendous amount of research on human motivation has been conducted, "there has been little systematic guidance for those who are trying to learn how to be more predictably effective in motivating their learners" (p. 3). He stated further, "The history of the study of motivation is long and deep. In fact, nobody knows how long it is" (p. 19).

Summary of the Literature

According to Kentnor (2015), "From the Postal Service, to spark transmitters, to television broadcasting, to the Internet and the Web, advances in communication technology have led to the changing landscape of education and the proliferation of distance education" (p. 30). The author also noted, "Online education is the fastest growing form of distance education and is valued at both traditional and non-traditional colleges and universities" (p. 30). The continued evolution of digital technologies has enabled the "online multilateral exchanges of visual, text, and audio within and outside of the learning community" (Clark & Mayer, 2011, p. 7). Such instruction that incorporates words and pictures is referred to as multimedia instruction (Mayer, 2001). In order to design effective multimedia instruction for this study, Clark and Mayer's (2011) multimedia principles were presented along with schema and cognitive load theories, which serve as a foundation for those principles.

Scenario-based learning goes by many names such as case-based, immersive, or whole-task learning (Clark, 2016) and despite an abundance of research into these teaching methods (Breakey et al., 2008; Carpenter et al., 2003; Hursen & Fasli, 2017; Landrigan, 2010; Lehmann et al., 2013; Lim et al., 2009; Loyens et al., 2015; Meldrum, 2011; Sorin, 2013; Yarnall et al., 2007; Yetik et al., 2012), it is still apparent that SBeL is not a cure-all for instruction (Clark, 2016). Nevertheless, Clark (2016) stated, “Learning in the context of authentic work-related situations and problems can make the relevance of the instruction more salient than in many traditional approaches” (p. 55).

Similarly, while there are a vast number of research studies on feedback resulting in mixed results (Shute, 2008), according to Nicol and Macfarlane-Dick (2006), “There is considerable research evidence to show that effective feedback leads to learning gains” (p. 204). Burkšaitienė (2011) suggested, “Effective feedback can increase student effort and motivation or engagement to reduce the discrepancy between what is understood and what is aimed to be understood, and can foster their learning” (p. 43). This author elaborated stating effective feedback must be, “clear and understandable, purposeful, motivating, compatible with students’ prior knowledge, providing links to the gaps in their learning that have to be closed” (p. 43). For the current study, the term elaborative feedback will be used to mean informing learners of the correctness of responses as well as providing additional information. For example, informing learners of an incorrect sperm cell count and providing additional information to help address the error. The term intrinsic feedback will be used to mean providing learners information regarding how responses relate to the scenario of reporting manual sperm count results. For example, informing the learners of an incorrect sperm cell count by describing the consequences of that count such as a couple needlessly spending thousands of dollars to conceive. The feedback for all treatment conditions will be delivered as part of the instructional interventions as on-screen text.

Elements of Keller's (2010) ARCS model of motivation were presented as a framework for evaluating learners' motivation. According to Keller, "Learner attention is necessary for both motivation and learning" (p. 76). Similarly, without relevance, students will not be motivated to learn (Keller, 2010). In addition, if learners do not have confidence in their ability to succeed, their motivation will likely suffer. Keller suggested, "To help students improve their confidence, provide corrective feedback that helps them see the causes of their mistakes and how to take corrective action" (p. 52). Finally, feelings of satisfaction and accomplishment can sustain learner motivation (Keller, 2010).

CHAPTER III

Methods

The goal of this study was to evaluate mean differences in student learning outcomes and motivation over time as a function of scenarios and two types of feedback as well as the interaction between these variables. This research employed a pair of 2×2 repeated measures ANOVAs with between group factors and random assignment. Participants' absolute cell counts were also analyzed to determine whether participants demonstrated mastery as indicated by correctly answering two out of three cell count questions to within ±1 standard deviation of a panel of experts. Because the pool of participants at the researcher's institution was relatively small, the researcher attempted to solicit participants from similar Medical Laboratory Science (MLS) and Medical Laboratory Technician (MLT) programs nationwide to obtain a larger sample. When those efforts were unsuccessful, the researcher opened participation to university students in general as well as instructional designers.

Research Questions

To examine the effects of scenarios and feedback type on student learning and motivation, the following research questions were examined:

1. Does the use of SBeL and feedback type have a significant effect over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest?
    a. Does the use of SBeL have a significant main effect over time on college students' ability to manually count sperm cells as measured by the

comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest?

b. Does feedback type have a significant main effect over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest?

c. Is there a significant interaction effect of the combination of SBeL and feedback type over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest?

2. Does the use of SBeL and feedback type have a significant effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's (2010) Instructional Materials Motivation Survey (IMMS) pretest, immediate posttest, and delayed posttest?

   a. Does the use of SBeL have a significant main effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's IMMS pretest, immediate posttest, and delayed posttest?

   b. Does feedback type have a significant main effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's IMMS pretest, immediate posttest, and delayed posttest?

c. Is there a significant interaction effect of the combination of SBeL and feedback type over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's IMMS pretest, immediate posttest, and delayed posttest?

Research Design

As shown in Table 1, two independent variables, scenarios and feedback, each with two levels (present/not present and elaborative/intrinsic respectively), were administered across four groups. The results of a researcher-developed performance pretest were used in an effort to account for individual differences between groups. The results of Keller's (2010) IMMS pretest were also used to account for individual differences in participants' motivation for learning the content. The results of the immediate and delayed researcher-developed performance posttests were compared to determine the mean differences, if any, in performance between groups over time. The mean difference in motivation between groups over time was examined by comparing the results of Keller's IMMS immediate and delayed posttests. Additionally, a demographic survey was administered to gather a more complete profile of the participants. The survey included gender, age, program type, and whether participants had previous experience with the content and online delivery method. The results of the demographic survey were only used for descriptive purposes.

Table 1

Research Design

R	S	O_1	O_2	X_1	O_3	O_4	O_5	O_6
R	S	O_1	O_2	X_2	O_3	O_4	O_5	O_6
R	S	O_1	O_2	X_3	O_3	O_4	O_5	O_6
R	S	O_1	O_2	X_4	O_3	O_4	O_5	O_6

R Random assignment.
S Demographic survey.
O_1 Motivation pretest, Keller's (2010) Instructional Materials Motivation Survey.
O_2 Researcher-developed content pretest.
X_1 Non-SBeL module with elaborative feedback.
X_2 Non-SBeL module with intrinsic feedback.
X_3 SBeL module with elaborative feedback.
X_4 SBeL module with intrinsic feedback.
O_3 Immediate motivation posttest, Keller's Instructional Materials Motivation Survey.
O_4 Immediate researcher-developed content posttest.
O_5 Delayed motivation posttest, Keller's Instructional Materials Motivation Survey.
O_6 Delayed researcher-developed content posttest.

As shown in Table 2, the two levels of feedback were elaborative and intrinsic and the two levels of the scenario variable were with or without a scenario. Therefore, there were four treatment conditions, (a) no scenario with elaborative feedback (sE), (b) no scenario with intrinsic feedback (sI), (c) a scenario with elaborative feedback (SE), and (d) a scenario with intrinsic feedback (SI).

Table 2

Treatment Conditions

	Elaborative Feedback	Intrinsic Feedback
No Scenario	No scenario, elaborative feedback (sE)	No scenario, intrinsic feedback (sI)
Scenario	Scenario, elaborative feedback (SE)	Scenario, intrinsic feedback (SI)

Note. Description of potential treatment conditions.

Participants were randomly assigned to the sE, sI, SE, or SI groups during the first week of the intervention. After being randomly assigned to one of the treatment groups, participants completed Keller's (2010) IMMS online to determine participants' confidence and motivation before starting the treatment. Participants then completed the researcher-developed online content pretest to determine prior knowledge of the content and ability to manually count sperm cells. Results of the content pretest were used to account for individual differences and therefore increase power. Participants completed the instructional intervention during the fourth week of the study. After completing the instructional intervention, participants completed Keller's IMMS as an immediate posttest to measure their confidence and overall motivation after the instruction. After completing the IMMS immediate posttest, participants also completed the immediate, researcher-developed, performance posttest. Participants completed Keller's IMMS and the delayed, researcher-developed, performance posttest during week seven to examine participants' longer-term motivation and learning. To address research question one, mean differences between group results on the immediate and delayed researcher-developed posttests were examined. Differences between the immediate and delayed IMMS posttests were analyzed to address research question two.

Treatment

Participants were randomly placed into one of four treatment conditions (sE, sI, SE, and SI). All treatment conditions addressed the concept of manually counting sperm cells and all treatments were delivered online. The content of the modules was created using the knowledge and real-world experience of the course instructor as well as input from professionals as subject matter experts (SMEs). The current instructor had previously taught this course six semesters. All treatments were delivered using web-based modules in the university's community learning management system (LMS); therefore, participants required an Internet connection to navigate through the instructional modules online. Students in the MLS program at the researcher's institution were required to have computer and Internet access and because the call for additional participants was delivered online, other participants outside of the university MLS program also had Internet capabilities. Access to this technology could limit the generalizability of the study results to such learners. An orientation familiarized participants to the treatment format and delivery method. The orientation was a demonstration of the delivery method and treatment, which was delivered as a pre-recorded step-through slideshow. Participants' knowledge of the technology used to deliver the instruction may have been different; therefore, the orientation served to minimize prior differences in experience with the delivery format. Participants were able to navigate the instructional modules at their own pace and the instructional modules during week 4 were available to participants for seven days.

All of the treatments were designed based on the existing text-based standard operating procedure (SOP) for manually counting sperm cells. The SOP outlines the principles behind the semen analysis procedure, its clinical significance, safety considerations, the specimen type, any special collection/preservation instructions, required supplies for performing the analysis, quality control standards, the procedure, reference intervals, and instructions for interpreting the results of the analysis.

The four instructional modules were differentiated based on the levels of the independent variables. The first module did not have a scenario, but did have elaborative feedback (sE). Therefore, the instruction was presented without framing it within the context of working in a medical laboratory. Instead, the instruction was presented as the steps necessary to manually count sperm cells and calculate an absolute total count. According to Shute (2008), elaborative feedback “provides information about particular responses or behaviors beyond their accuracy” (p. 157). Shute explained that the “Elaboration can (a) address the topic, (b) address the response, (c) discuss the particular error(s), (d) provide worked examples, or (e) give gentle guidance” (158). For the purposes of this study, elaborative feedback addressed the participant’s response and provided guidance. If participants answered correctly, the elaborative feedback indicated the response was correct and included a summary of the procedural steps. If participants answered incorrectly, the feedback indicated the response was incorrect and provided an explanation for why the answer was incorrect as wells as guidance to help the participant answer correctly (the summary of the procedural steps). Because the main interest in this study was the efficacy of a scenario-based instructional approach with intrinsic feedback, the sE treatment served as the control group for this study.

The second instructional module did not include a scenario, but provided intrinsic feedback (sI). The instruction followed the same pattern as the sE treatment; however, the feedback was related to the consequences of the participant’s answers. The consequences related to being deemed proficient as a medical laboratory employee and keeping the dream job as well as two types of errors, which were over-counting and under-counting sperm cells. In the case of over-counting, a patient would incorrectly be informed of a fertile sperm count. In such a case, a patient would believe he could conceive, but in reality the sperm count is too low. Affective domain consequences, such as disappointment or feelings of inadequacy, could result after multiple failed attempts at conception. The couple may also fail to explore other options such as female infertility. Additionally, continued attempts at conception could take time, which could have the effect of advancing the female partner past an age where conception is advised or her fertility is diminished.

Alternatively, in the case of under-counting, a patient would incorrectly be informed that he did not have an adequate sperm count when in reality he did, which could also result in negative affective domain consequences. Believing a diagnosis of infertility could affect a patient's self-esteem and lead to feelings of disappointment and inadequacy. Another consequence of a false low count could be financial. If the male is thought to be infertile, a couple may incur additional expenses examining female fertility issues or investing in alternative fertility treatments.

The third module had a scenario and elaborative feedback (SE). The scenario centered on participants being employed at a theoretical lab and receiving an order to examine a patient specimen and report the results. In the medical laboratory, this order is known as a patient requisition. In the SBeL treatment, participants were introduced to their new supervisor who had them perform sperm cell counts as part of their training and competency to perform semen analysis in their new workplace. As with the first treatment condition, the elaborative feedback in this treatment confirmed the accuracy of the participant's report and addressed errors in order to help guide participants to the correct answer.

The fourth treatment combined a scenario with intrinsic feedback (SI). The scenario was again based on a patient requisition and participants reported the results of the cell count to their new supervisor. The intrinsic feedback not only confirmed participants' responses as correct or incorrect, but also provided participants with the consequential results of their report based upon the scenario. In addition to being able to keep the scenario dream job, the consequences related to a patient's belief in his ability or inability to conceive a child. The consequences of both false positive and false negative results could be affective as well as financial.

Participants and Sampling

The majority of the participants for this study were undergraduate and graduate students majoring in a Medical Laboratory Science (MLS) program at a medium-sized university in the Intermountain West. These students were from the researcher's institution and participated for credit in an eight-week, online Urinalysis and Body Fluids course. During the fall 2017 and fall 2018 semesters, 48 and 49 students completed the Urinalysis and Body Fluids course (respectively) and the fall 2019 semester had similar enrollment with 48 students. Comparable programs are also delivered at universities across the United Sates and the researcher attempted to solicit additional participants from those institutions to increase the sample size for this study.

To obtain participants from other institutions, a message was posted on a listserv requesting any interested parties contact the researcher. The researcher received correspondence from faculty at ten other institutions from Washington D.C., Texas, South Dakota, Ohio, Missouri, Mississippi, Kentucky, Iowa, and Colorado. Outside faculty would have been asked to submit written letters of support expressing their willingness to participate and to implement the instruction in the same manner as the researcher's institution. Program numbers and demographics would also have been collected from faculty at each participating institution to provide a precise description of those participants; however, because no outside institutions participated in the study, no such information was gathered.

After being unable to acquire MLS and MLT participants from outside institutions, the researcher placed calls for additional participants through university student memos, student organization advisors, and various social media channels. Through these means, the researcher was able to solicit non-MLS/MLT participants that included students, faculty, and instructional designers. These participants completed the study in three additional treatment iterations over the course of the spring 2020 semester. All participants were pooled to compare results across the treatment levels.

The researcher obtained Human Subjects approval from the primary institution. Prior to the treatment, the researcher obtained informed consent from all participants as part of the introductory activities, which occurred during the first week of each treatment. While participants at the researcher's institution were required to complete all of the study materials as part of a course grade, non-MLS/MLT participants completed the materials on a strictly voluntary basis; however, the researcher added Amazon gift card drawings as an incentive for participating. All participants had the option to exclude their data from the research study at any time.

Instruments

Multiple instruments were used for this study. Performance and content knowledge were assessed using a researcher-developed instrument, while motivation and confidence were measured using Keller's (2010) IMMS instrument (used with permission, see Appendix B). All instruments were delivered online using an LMS at the researcher's institution. Each of these instruments is addressed next.

The first instrument was designed by the researcher to assess both participants' content knowledge and ability to perform manual sperm cell counts. This content and performance assessment was developed using MLS faculty and professionals as SMEs. The pretest contained 25 questions while the immediate and delayed instruments contained 27 questions. The reason for this difference was that the pretest only included one absolute cell count question rather than the three absolute count questions in the immediate and delayed posttests to avoid pre-teaching the skill. The content knowledge aspect of the instrument was designed in conjunction with the instructional interventions and was evaluated using the Delphi method with MLS faculty and professionals to determine usability, validity, and reliability. Content validity of the assessments was determined by SMEs who examined the objectives and instruction to determine that the assessments covered the objectives without over-emphasis or omissions. The performance aspect of the instrument was designed similarly to proficiency procedures participants will encounter if they are hired at a lab where cell counts are conducted. According to the Centers for Medicare & Medicaid Services website (2018), The Clinical Laboratory Improvement Amendments (CLIA) passed in 1988 established quality standards for proficiency testing. As part of the proficiency testing protocol, employees must analyze samples as though they are real patient samples to demonstrate their ability to accurately examine and report the tests performed in the laboratory. Depending on the number of employees in a particular lab, a certain employee may not be assigned to perform the proficiency test. Although it is not required that each person working in the lab be able to complete the test, it is best practice to rotate individuals so all employees participate in the proficiency testing process over the course of one year. A medical laboratory director must subscribe to a proficiency program; however, they have the option to choose from different vendors. The College of American Pathology (CAP) is among the first organizations to offer proficiency packages, which are world-renowned (S. Galindo, personal communication, April 17, 2019). The CAP product is used locally and is one of the most common nationwide; therefore, the current study utilized this model. The labs that subscribe to the CAP semen analysis proficiency package receive two challenges twice per year to demonstrate proficiency. The CAP semen analysis package includes digital video to assess motility and count as well as digital slide images to assess viability and morphology. Although not all lab scientists perform semen analysis, performing a manual cell count is considered a fundamental skill for lab scientists and proficiency counting white blood cells, red blood cells, and platelets is required as part of CLIA (S. Galindo, personal communication, February 17, 2017). A clinician will interpret many parameters, including morphology, motility, viability, and count, to provide a fertility diagnosis to the patient; however, including all of the parameters in a semen analysis, such as assessing motility and morphology, would likely have confounded the results and made it difficult to determine where participants struggled. Instead, the morphology and percent motility were included in the instruction and participants only reported the sperm cell count. To demonstrate proficiency for the current study, participants were required to perform three sperm cell counts each within plus or minus one standard deviation (±1 SD) of the mean of at least five expert counts. The knowledge and skill assessment was used as a pretest, immediate posttest, and delayed posttest, each having different absolute counts. The immediate and delayed posttest questions were delivered to all participants in random order.

The second instrument for this study was used to gauge participants' confidence in their ability to manually count sperm cells and their motivation for doing so. Keller (2010) designed the IMMS, "to measure learner reactions to self-directed instructional materials" (p. 277). The IMMS consists of 36 items in four subscales (attention, relevance, confidence, and satisfaction). The IMMS was administered three times, as a pre-survey, immediate post-survey, and delayed post-survey.

The third instrument was a demographic survey to collect participants' gender, age, program type, and whether they had previous experience with the content and online delivery method. The data from this instrument were not used as variables in this study. Instead, the results of this survey were used for descriptive purposes to provide a more detailed view of the sample.

Data Collection and Analysis

After obtaining approval from Human Subjects and informed consent from the participants, the demographic survey, Keller's (2010) IMMS pretest, and the researcher-developed performance pretest were administered during the first week of each treatment in that order. The instructional intervention was delivered during week four and was followed by the immediate IMMS posttest and the researcher-developed performance posttest. The immediate IMMS posttest was administered prior to the immediate performance posttest to avoid the potential influence of perceived poor content performance on participant confidence and overall motivation. Participants completed a delayed IMMS posttest and delayed, researcher-developed, performance posttest during week seven.

The performance and Keller's (2010) IMMS results were downloaded from the LMS as Excel spreadsheets. Data from the spreadsheets were entered into IBM SPSS (Version 26.0) for statistical analyses. The differences between treatment groups over time were analyzed using repeated measures tests of analysis of variance (ANOVA). Although participants can experience fatigue with multiple measurements, the repeated measures ANOVA was chosen because it can minimize problems related to individual differences, it requires relatively few participants, and it has increased power over independent-measures tests (Gravetter & Wallnau, 2007). Participants' absolute cell count results were also compared to expert counts to determine mastery. Three absolute cell counts were included on the immediate and delayed posttests. For answers to be considered correct, participant counts were required to be within plus or minus one standard deviation (±1 SD) of the mean of at least five expert counts. Mastery was met if participants correctly answered at least two out of the three counts, whereas mastery was not met if participants correctly answered one or fewer count questions.

To provide a more complete description of the sample, learner demographics were collected and reported. In addition, because participants in the second through fourth iterations could have been from MLS or MLT programs, the researcher collected and reported that information, but did not plan to compare the results across these two groups. Because no outside MLS or MLT programs joined the study, the only MLS/MLT participants were from the researcher's institution. The vast majority of remaining participants were from the general population at the researcher's university. The primary difference between these programs is the length of the program, which results in different certification requirements. An MLS program requires four years while an MLT program takes two years to complete, and therefore the MLS professional has more advanced training.

Development of the Treatment

Morrison, Ross, Kalman, and Kemp's (2011) systematic instructional design approach (Kemp model) was used to design, implement, and evaluate the treatments for this study. The Kemp model is an integration of four fundamental planning elements, the learners, objectives, methods, and evaluation. The Kemp model expands those elements to include instructional problems, learner and context, task analysis, instructional objectives, content sequencing, instructional strategies, designing the message, developing the instruction, evaluation instruments, planning and project management, support services, formative evaluation and revision, implementation, summative evaluation, and confirmative evaluation. Each of these elements will be discussed further in the following sections.

Instructional problems. Before designing instruction, Morrison et al. (2011) recommended conducting a needs analysis to determine whether instruction is necessary. The analysis begins with the identification of an instructional need. For the proposed study, the researcher worked closely with an MLS faculty member to determine the instructional need for manual cell counting. Because manual sperm cell counts are difficult to perform in a classroom laboratory and because not all of the students in the program attend on campus, there was a need to deliver computer-based instruction. The instructional goal of the treatment was to provide learners with the knowledge, confidence, and motivation to manually count sperm cells (which should transfer to other cell types) and the treatment was designed to achieve these goals.

Learner and contextual analysis. Certain characteristics of the learners and the learning environment will affect how instruction is both designed and delivered (Morrison et al., 2011). Participants at the researcher's institution were enrolled in a Body Fluids and Urinalysis course. This course is taken during the first semester of the MLS program at the researcher's institution. Both undergraduate and graduate degree programs are offered at the researcher's institution and students would have completed the university's general education goals prior to admission to either program. Additionally, a 2.5 grade point average is the minimum for acceptance into the program. The instructor of this course at the researcher's institution assumes the learners enter the course with no previous hemocytometer experience; however, it was assumed the participants had some prior experience with microscopes. On average, there were 32 MLS program graduates each semester between fall 2009 and spring 2018. Undergraduate female students represented the majority of program graduates. From fall 2009 through spring 2018, 96% of students were undergraduates and 66% were female. Data on age, race, and ethnicity were either unavailable or unknown; therefore, non-traditional students and racial breakdown were not considered for this study.

The contextual analysis includes examining how and where the instruction will take place. According to Morrison et al. (2011), "context influences every aspect of the learning experience" (p. 65). Additionally, providing context "makes the content concrete and realistic and helps the student understand . . . how it can be applied on the job" (Morrison et al., 2011, p. 65). For the purposes of this study, two of the four treatments included real-world context using scenarios. Context also relates to the physical environment in which instruction is delivered. As Morrison et al. illustrated, students learning in a classroom near a break room may be disrupted by noise or other distractions. Participants in the proposed study chose their own location to complete the instruction; therefore, external environmental factors were difficult to control.

Task analysis. Morrison et al. (2011) "refer to *task analysis* as the collection of procedures for defining the content of an instructional unit" (p. 78, emphasis in original). The instruction for the proposed study focused on teaching a procedure; therefore, a procedural analysis was used. This analysis is used to divide the task into the steps required to complete the procedure. The main task of the instructional intervention for this study will be to provide students with the knowledge, skills, confidence, and motivation to manually count sperm cells using a hemocytometer. As such, participants needed to know how to (a) identify sperm cells, (b) calculate volume given the dimensions of a three-dimensional counting chamber, and (c) convert microliters (μL) to milliliters (mL). Additionally, participants needed to know some scientific notation and terminology as well as the required equipment and reagents.

Instructional objectives. Clear and measurable objectives are required to determine what to include in the instruction and how to evaluate learning outcomes. The following learning objectives were used to design and evaluate the treatments for this study:

1. Given instruction regarding the use of semen testing, the participant will be able to identify the clinical reasons for performing semen testing.

2. Given instruction regarding how to prepare a semen sample for analysis, the participant will be able to recall the steps required to charge both sides of a clean hemocytometer for semen analysis and apply quality control specifications to determine if the sample is evenly distributed between both sides.
3. Given an image of sperm cells in a hemocytometer, the participant will be able to identify and count the number of sperm cells on both sides of the hemocytometer.
4. Given the sperm count from the two sides of the hemocytometer chamber, the participant will be able to convert a fractional cubic volume to a cubic millimeter volume.
5. Given the initial dilution of the sample and total volume of semen, the participant will be able to determine and report the number of sperm per mL and the absolute number of sperm in the entire specimen to within plus or minus one standard deviation (±1 SD) of the mean of at least five expert counts.

Content sequencing. According to Morrison et al. (2011), "Sequencing is the efficient ordering of content in such a way as to help the learner achieve the objectives" (p. 136). Because the instruction for this study concerned teaching a procedure, the content was sequenced temporally according to the procedure. Galbraith (2011) found that students who received instructor-sequenced content scored higher on an achievement test than students who chose their own instructional sequence. Galbraith asserted that, "given the experience that instructors possess, they become skilled at sequencing course content in an optimal, or near optimal, order even though other possible sequential orders exist" (p. 74). For that reason, the faculty member acting as a SME attempted optimal content sequencing.

Instructional strategies. According to Morrison et al. (2011), there are two levels of decisions for designing instruction. The first is the delivery strategy, which is how the instruction will be presented to learners. The second level is the instructional strategy, which encompasses the sequences and methods of instruction. For the proposed study, the interventions were delivered as multimedia web-based modules. The instructional strategy for all treatments included a worked example demonstration with two practice examples for participants to complete. The instructional modules either had scenarios or no scenarios, whereas the feedback was only included with the practice activities. The sE treatment practice used elaborative feedback without scenarios. The sI treatment practice used intrinsic feedback without scenarios. The SE treatment practice had elaborative feedback with scenarios and the SI treatment practice had intrinsic feedback with scenarios.

Designing the message. The four treatment conditions were developed using a combination of text, graphics, and video in accordance with Clark and Mayer's (2011) cognitive theory of multimedia learning (CTML) principles discussed in Chapter II. Mayer (2014c) categorized these principles into those aimed at reducing extraneous processing, managing essential processing, and fostering generative processing. The instruction for the current study incorporated certain multimedia principles to meet these goals and is discussed next according to those aims.

Mayer (2014c) asserted the coherence, signaling, and redundancy principles as well as the spatial and temporal contiguity principles are intended to reduce extraneous processing. As mentioned previously, the coherence principle aims to reduce extraneous processing by limiting extraneous audio, graphics, and words, which can increase extraneous cognitive load and distract learners from the learning objectives at hand (Clark & Mayer, 2011). Van Gog (2014) asserted that the signaling principle is important during the selecting phase of the CTML. This author stated, "To enable learning from multimedia materials, information needs to be attended to in order to be available for processing in working memory" (p. 263). Furthermore, Mayer and Fiorella (2014) stated that the signaling principle reduces extraneous cognitive load by directing the learner's attention to the essential material, "thereby enabling the learner to ignore extraneous material and use more available cognitive capacity to process essential material" (p. 292). Similarly, Clark and Mayer (2011) asserted the redundancy principle reduces cognitive load by limiting redundant on-screen text. According to Mayer (2014c), the spatial and temporal contiguity principles are intended to reduce extraneous processing by placing printed words in close proximity to corresponding graphics and presenting narration simultaneously with corresponding graphics (respectively).

Clark and Mayer's (2011) principles for reducing extraneous cognitive load that will not be incorporated in this study are the temporal contiguity and redundancy principles. The temporal contiguity and redundancy principles were held constant for all treatments because the instructional interventions did not include narration.

Clark and Mayer's (2011) principles for reducing extraneous cognitive load that were included in this study are the spatial contiguity and coherence principles. By keeping text responses and feedback together, the instruction for this study met the spatial contiguity principle. Furthermore, this study adhered to the coherence principle by excluding extraneous background music and limiting extraneous text and graphics. As mentioned previously, according to Mayer (2014c), signaling refers to drawing the learners’ attention to essential material and its organization, which was accomplished in the current study by using headings and pointer words to indicate the order of the procedural steps. Van Gog (2014) also mentioned, “signaling not only may prevent extraneous load by preventing processing of less relevant information, but may simultaneously foster germane load, by facilitating the organization or integration of essential material” (p. 264).

Mayer (2014c) suggested incorporating the segmenting, pretraining, and modality principles to manage essential processing. Essential processing is the cognitive processing inherent in the content. According to Clark and Mayer (2011), the modality principle refers to putting “words in spoken form rather than printed form whenever the graphic (animation, video, or series of static frames) is the focus of the words and both are presented simultaneously” (p. 117). Mayer (2014c) stated, “the modality principle allows learners to off-load some of the processing in the visual channel (i.e., the printed captions) onto the verbal channel, thereby freeing more capacity in the visual channel for processing the animation” (p. 65); however, as mentioned previously, Clark and Mayer asserted this principle is more important when the instruction is fast-paced, the graphics are complex, and the words are familiar. Conversely, according to Clark and Mayer, this principle is less important when learners can control the pacing of the instruction. The instruction for the current study covered manual cell counts, did not consist of complex graphics or animations, and did not include narration. Furthermore, learners had control of the instructional pacing, which, according to Clark (2014), reduces the negative impact of written text describing on-screen visuals. Therefore, the absence of the modality principle was held constant across all treatments.

Use of the segmenting principle is thought to manage essential processing by presenting instructional material to learners in manageable parts that do not overwhelm working memory (Mayer, 2014c). In an effort to help the learners manage essential processing, the instruction for the current study was segmented using the natural breaking points involved in the cell counting procedure. In the event the natural breaking points appeared to have too many concepts or facts presented to learners at one time, the content was examined to determine segments that were more manageable. Additionally, learners controlled the pace of the lesson.

Clark and Mayer (2011) recommended applying the pretraining principle to provide learners with key names and concepts, which allows learners to focus on essential processing rather than the pretrained definitions and foundational concepts. For the current study, pretraining was used in an attempt to provide learners with a consistent, minimum prerequisite skill level by familiarizing all learners with the treatment delivery method. This pretraining was delivered as an orientation in week one of the study. The orientation summarized how participants would navigate the treatment and the tools they would use during the treatment. Following the week one orientation, the week 4 instructional interventions were differentiated based on the treatment conditions as discussed previously.

Mayer (2014c) suggested incorporating the voice, embodiment, image, and personalization principles to foster generative processing. According to Mayer (2014a), these principles are based on social cues in multimedia learning and aim to create a social presence for the learner, which may increase learner engagement, and thus generative processing, during the instruction. According to the image principle, including the narrator's image on screen does not necessarily promote deeper learning. The embodiment principle also pertains to on-screen agents, which Mayer (2014c) suggested should have human-like gestures. The voice principle refers to using a human instead of machine voice for narration. These first three principles were held constant across the treatments by not including narration, a narrator's image, or a pedagogical agent; however, the current study did incorporate the personalization principle.

According to Mayer (2014a), the personalization principle refers to using words in a conversational rather than formal style. To meet the personalization principle, the on-screen text for this study's instructional interventions was written in a conversational tone.

Developing the instruction. The instruction for this study was designed as self-paced e-learning modules. To design quality self-paced learning, Morrison et al. (2011) suggested carefully designing the instruction to meet specific objectives, carefully selecting the activities and resources used to meet the objectives, checking student mastery before proceeding to the next step, and including feedback. As discussed previously, each of the four treatment interventions (sE, sI, SE, and SI) were designed to meet the Morrison et al. criteria for quality self-paced learning by using either elaborative or intrinsic feedback. The researcher and the course instructor created resources and designed the treatments specifically to address the learning objectives outlined in the instructional objectives section. Additionally, because minimally-guided instruction is thought to be less effective (Kirschner, Sweller, & Clark, 2006), the instruction was designed with what was thought to be an appropriate amount of guidance to minimize the flounder effect (Clark, 2013).

For each iteration of this study, the researcher created courses within a community LMS available through the researcher's institution. The courses consisted of seven weeks with introductory materials and pretests in week one, the instructional interventions and immediate posttests in week four, and the delayed posttests in week seven. For the instructional interventions, the researcher created four learning objects (LOs) centered on performing manual sperm cell counts. The LOs were created using the Lesson activity in the LMS that consists of multiple web pages learners navigated using a Lesson menu (table of contents) as shown in Figure 4.

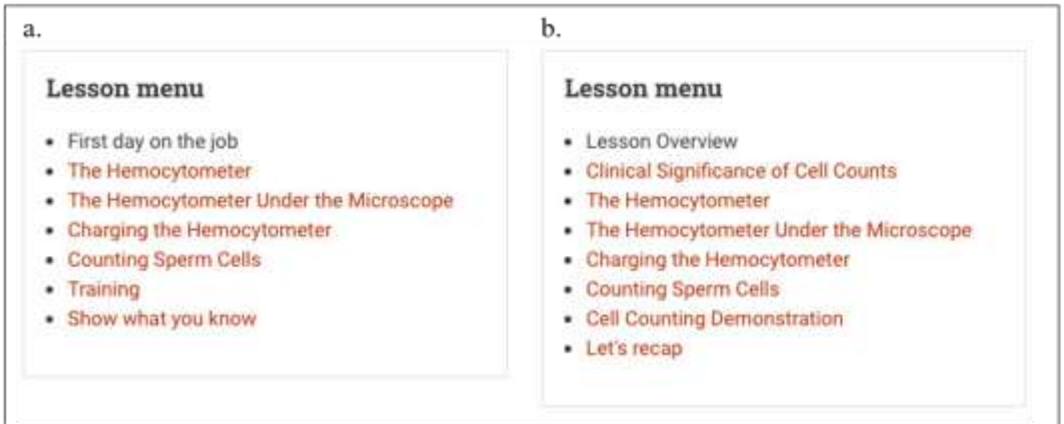


Figure 4. Lesson table of contents. The table of contents for the (a) scenario treatment and (b) no scenario treatment. Note, the no scenario Clinical Significance of Cell Counts information was included on the scenario-based First day on the job page.

Alternatively, the Lessons included navigation buttons. Figure 5 shows the navigation buttons on one of the scenario-based Lesson pages, but the navigation buttons were the same for scenario and no-scenario treatments.

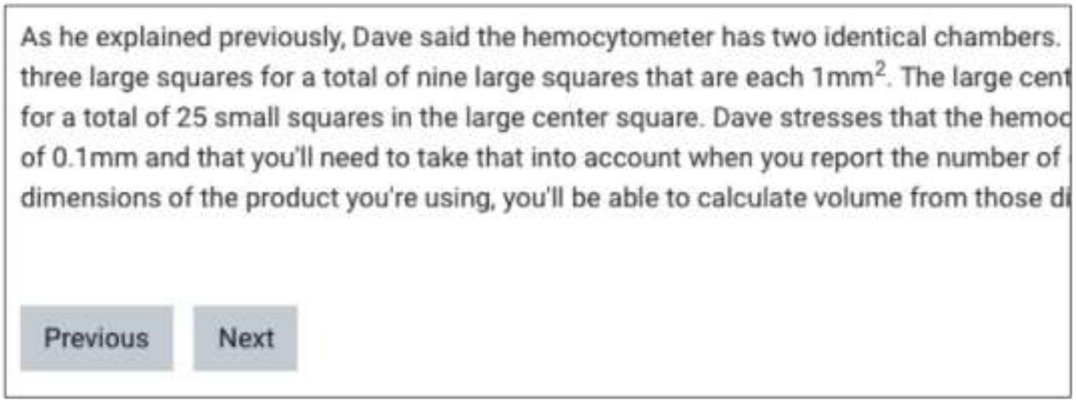


Figure 5. Lesson navigation buttons.

Scenarios were used to present the material in two LOs while the other two LOs did not include scenarios in the Lesson pages. For example, the scenario-based LOs started with, "Congratulations! You've just landed your first dream job in a medical laboratory. One of the first tests that have come in is a semen analysis…" whereas the two LOs without scenarios began, "This lesson is designed to show you how to manually count sperm cells using a standard hemocytometer. A hemocytometer is a specialized microscope slide." Similarly, the remaining Lesson pages contained the same information, but the narratives varied by scenario level (present or not present). Refer to Appendix C for a detailed comparison of the Lesson pages.

The Hemocytometer page provided learners an overview of the hemocytometer slide and its parts. The *Hemocytometer Under the Microscope* page presented learners with an explanation of the grid used for cell counts. The *Charging the Hemocytometer* page provided learners with the procedure for adding samples to the hemocytometer chambers and included the same demonstration video for both scenario and no scenario treatments. The *Training* (scenario) and *Cell Counting Demonstration* (no scenario) pages of the Lessons provided students with worked examples, which were differentiated based on scenario level and are compared in Appendix D. After completing the worked examples, students were given the opportunity to check their ability to put the procedural steps in order on the *Show what you know* (scenario) and *Let's recap* (no scenario) pages.

Following this page, learners were presented with two practice activities that were not included in the table of contents to avoid learners navigating to these activities without completing the instruction. The practice activities were the same for scenario and no scenario treatments; however, the presentation again differed based on treatment. See Appendix C for a comparison of the practice pages based on scenario level and Appendix E for screenshots of the practice activity. The feedback presented to learners as they completed these practice activities differed based on feedback type.

Two of the LOs provided learners with elaborative feedback while the other two provided learners with intrinsic feedback. For example, the intrinsic feedback for one incorrect answer stated, "Unfortunately, your count was significantly less than Dave's calculation of 857,500,000 as well as below the reference range. The consequences of reporting such an undercount include the clinician informing a couple that they're unlikely to conceive naturally based on erroneous information, which may cause them undue stress. Additionally, the clinician may also suggest more costly and extensive fertility testing and procedures…"

Alternatively, the elaborative feedback for the same question stated, "Incorrect, the correct answer is 857,500,000. It looks like you may have calculated incorrectly. We only counted 8 small squares in this practice so rather than dividing by 10 small squares, you would divide by 8. Additionally, the dilution factor was 4 and the total ejaculate volume was 2.0 mL…" Both feedback types included a summary of the correct steps as well as the values that should have been obtained as shown in Figure 6. A comparison of the elaborative and intrinsic feedback can be found in Appendix F.

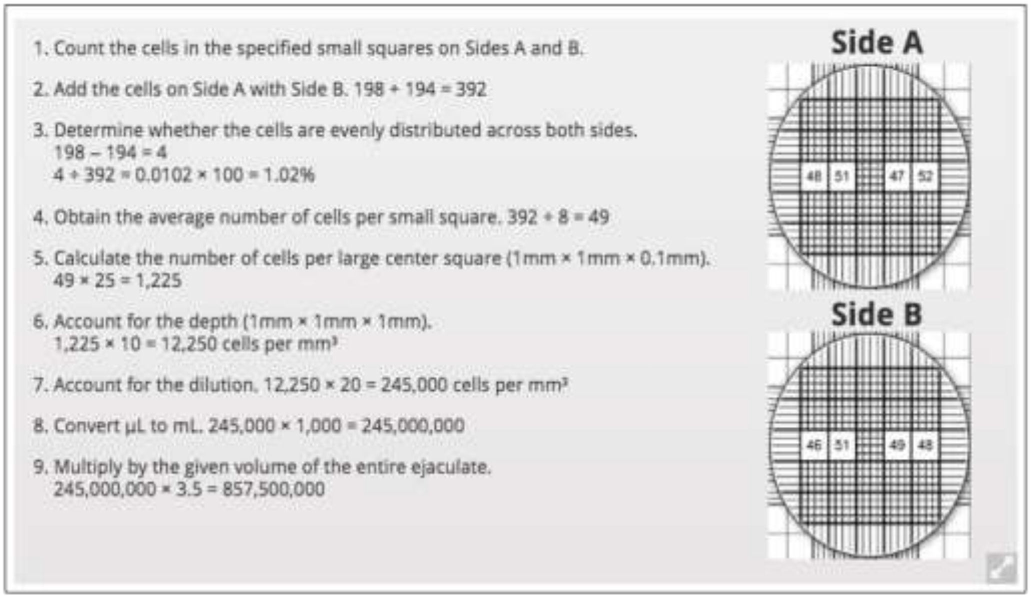


Figure 6. Feedback procedural summary. This summary was provided as part of elaborative and intrinsic feedback.

A small-scale pilot (N = 5) was conducted with MLS and MLT faculty, recent program graduates, and practitioners to test the instruction for usability and functionality. Of the 16 people who expressed interest in testing the LOs, 10 self-enrolled in the test course, and five accessed the materials. One test participant stated she did not find any typos or misinformation, but notified the researcher that the hemocytometer charging demonstration video did not work in the Safari web browser on an Apple computer. The researcher embedded the video directly in the lesson activity rather than using YouTube to resolve this issue. This participant also stated, "Great lesson plan! My favorite was the [scenario with intrinsic feedback] version…Can your next lesson plan be on CSF cell counts and other body fluids?" Another pilot participant stated, "The manual cell count is really awesome!"

The same faculty, program graduates, and experts were also asked to vet the knowledge instruments for content validity. One student tester expressed confusion by the difference between microliter (µL) and cubic millimeter (mm^3) on the pretest; however, these two terms were both used because they are equivalent measures of volume and this information would be included in the instruction. Additionally, an MLS faculty member stated, "There seemed to be quite a few questions regarding the comparison of the sides [of the hemocytometer]. I am wondering if you might not need that many?" However, the comparison questions were true/false questions, which meant students had a 50/50 chance of guessing correctly; therefore, the researcher and SME kept the ten comparison questions to ensure enough questions to represent the learners' ability to determine even distribution of cells across both sides of the hemocytometer. After receiving this explanation, the faculty member replied "that makes sense about the additional questions." She also stated, "I thought the questions for the actual chamber counting were good."

It was also recommended to inform learners to scroll down to see the answer field below the hemocytometer chamber in case the text field was out of view on their screens when taking the knowledge tests. Lastly, the question was asked, "Do the instructions emphasize that the activity needs to be completed in one sitting? I know that the corporate training modules always seem to allow for coming back later." Therefore, a statement was added to the instructions of each instrument informing learners to complete the instrument in a single attempt. A statement was also added to inform participants to select an answer for each question before moving to the next question as they would not be able to go back.

Additionally, ID experts (N = 2) were invited to examine the instruction for usability, functionality, and its adherence to the multimedia principles. One faculty inquired about whether "in the square shown" needed clarification in pretest questions; however, the researcher and SME wanted to determine whether learners were familiar with the boundaries of the hemocytometer grid used for counting sperm cells. This information was also presented in the instructional interventions. For these reasons, the question wording was not edited. The ID experts were asked to note the adherence of the LOs to the multimedia principles as well as any other necessary improvements and because no such comments were received, the LOs were considered to be acceptable.

It was also noted that a drag-and-drop practice question in the hemocytometer charging demonstration video did not have the correct point value, which the researcher fixed. Furthermore, the ID expert questioned the meaning of the 1:20 dilution in the instruction. After consulting with the SME, it was determined that in the case of dilutions, the colon is not quite the same as a true ratio. The SME confirmed that it is one in 20 rather than 1 to 20.

Evaluation instruments. Two evaluation instruments were used for this study. The first instrument was Keller's (2010) IMMS. The IMMS has four subscales (attention, relevance, confidence, and satisfaction). According to Keller, the attention subscale consists of 12 items (α = .89), the relevance subscale consists of nine items (α = .81), the confidence subscale consists of nine items (α = .90), the satisfaction subscale consists of six items (α = .92), and the IMMS as a whole has been found to be highly reliable (36 items; α = .96). The IMMS has also been found to be a valid instrument to measure situation-specific motivation (Keller, 2010). The second instrument was a researcher-developed, knowledge and performance assessment to evaluate participants' content mastery and ability to manually count sperm cells. This assessment required participants to complete the steps and report the quantity for a manual sperm cell count. The Delphi method was used with SMEs to determine the content validity of the instrument.

Planning and project management. According to Morrison et al. (2011), the complexity of an instructional design project will determine the effort needed for planning and project management. These authors explained that planning and project management are critical for developing and managing a project's schedule and budget. The current study did not have a budget, but the instructional interventions were needed prior to the start of the fall 2019 semester. The researcher was responsible for meeting this deadline.

Support services. Depending on the size of an instructional design project, there may be several support resources involved including graphic designers and programmers among others (Morrison et al., 2011). The treatments for the current study were small enough in scale that the researcher served as the main support person for development and production of the instruction.

Formative evaluation and revision. The formative evaluation process for this study started with the identified need for manual sperm cell count instruction. Once a prototype of the instruction was created it was also evaluated by professionals in the clinical laboratory to determine whether the instruction was aligned to the stated objectives. The instruction was revised as deemed necessary by the expert review.

Implementation. The researcher worked with faculty at the researcher's institution to ensure the instruction was implemented properly. Implementation instructions were not necessary for faculty at outside institutions because the researcher administered all treatments to outside participants to ensure the treatments were administered consistently.

Summative evaluation. According to Morrison et al. (2011), "Summative evaluation is directed toward measuring the degree to which the major outcomes are attained by the end of the course" (p. 275). For the purposes of this study, summative evaluation refers to measuring student-learning outcomes after the instructional treatments. Although summative evaluation can also refer to evaluating the instructional design product rather than participants' learning outcomes (Morrison et al., 2011), that aspect will not be included in this study.

Confirmative evaluation. Because this study was conducted as part of the requirements for a doctorate, it may be difficult to determine whether the instruction remains applicable over time; however, the researcher is likely to make recommendations for confirmative evaluation based on the results of this study.

Chapter IV

Results

The purpose of this study was to examine the effects of scenario-based e-learning (SBeL) for teaching manual cell counts as well as the effect of two different feedback types on learner performance and motivation. Two 2×2 repeated measures analyses of variance (ANOVAs) with between-group factors and random assignment were used to examine these effects in addition to a comparison of participants' absolute cell counts to expert counts using a chi-square test of independence.

Sample

The participants for this study were gathered from the Medical Laboratory Science (MLS) program at a medium-sized university in the Intermountain West. Because comparable programs are also delivered at universities across the United States, the researcher attempted to solicit additional participants from those institutions to increase the sample size for this study. After those efforts proved unsuccessful, participants were invited from other programs within and outside the researcher's university to attempt to gather enough participants for inferential analysis.

Because the bulk of the sample (95%, 63 of the 66 participants that completed all of the training and assessments) was from the researcher's institution, those demographic data are presented here. Total enrollment at the researcher's institution for the 2019-2020 academic year was 12,425 with 10,365 undergraduate students and 2,060 graduate students. As shown in Table 3, of the 12,425 enrolled students, approximately 42% (5,280) were men and 57% (7,127) were women and 18 students did not report gender. Also shown in Table 3, reported ethnicities included 73% (9,105) White, 12% (1,430)

Hispanic, 2% (207) Asian, 1% (177) American Indian or Alaska Native, 1% (146) African American, and 0.26% (32) Native Hawaiian or Pacific Islander. In addition, 5% (634) reported unknown race or ethnicity, 3% (339) reported two or more races, and 3% (355) were reported as nonresident aliens.

Table 3

Enrollment Numbers for Race/Ethnicity and Gender.

Race/ethnicity	Enrollment	Percent
White (non-Hispanic)	9,105	73.28
Hispanic	1,430	11.51
Unknown	634	5.10
Nonresident aliens	355	2.86
Two or more races (non-Hispanic)	339	2.73
Asian	207	1.66
American Indian or Alaska Native	177	1.42
African American	146	1.18
Native Hawaiian or Pacific Islander	32	0.26
Gender	Enrollment	Percent
Women	7,127	57.36
Men	5,280	42.49
Gender not reported	18	0.14
Total	12,425	100.00

Note. Data are listed from highest to lowest enrollment.

There were a total of 84 participants who enrolled in one of the four courses that were delivered as part of this study. However, 18 students opted out or did not complete all of the research materials for a total of 66 participants (Table 4).

As mentioned previously, the study materials were delivered four times. The first iteration included 48 students in the MLS program at the researcher's institution. Of those participants, two opted out and one did not complete, which yielded 45 participants for the first iteration and a 98% completion rate. The participant who did not complete the study finished all of the research materials except Keller's (2010) delayed Instructional Materials Motivation Survey (IMMS) and the delayed knowledge posttest.

To avoid obtaining potentially biased data arising from an incomplete data set, the data were not included for participants who did not complete all of the study materials.

Participants for the second iteration were generated from an announcement posted to the online student bulletin board at the researcher's university. Additionally, the researcher contacted biology, microbiology, anatomy and physiology, and College of Education faculty at the researcher's institution requesting that they share the call for participants with their students. Seventeen participants started the second iteration; however, one student who opted out did not complete all of the research materials in addition to seven other participants who did not finish. Therefore, the second iteration had a 56% completion rate. Of the students who did not complete the study, four quit participating after the week one activities, one failed to complete the knowledge pretest, and one failed to complete the delayed knowledge posttest.

To gather participants for the third and fourth iterations, the researcher posted a second student bulletin board notice and reached out to faculty and professionals outside the researcher's university. The call for research participants was posted in an instructional design organization newsletter as well as on personal and program social media accounts. The researcher also contacted faculty advisors of all student organizations at the researcher's university and asked them share the call for participants with students in those organizations. The third iteration started with 14 participants and 11 finished for a 78% completion rate while the fourth iteration started with five participants and one finished (20% completion rate). Two of the third iteration participants who did not complete the study failed to participate after the first week and one participant only completed the consent form and orientation. Similarly, one participant in the fourth iteration completed the consent form and orientation while the other three quit participating after completing the week one materials.

Table 4

Participant Numbers by Iteration

Iteration	Started	Opted Out	Non-completed	Completed
One	48	2	1	45
Two	17	1	7	9
Three	14	0	3	11
Four	5	0	4	1
Total	84	3	14	66

Note. Participants who opted out were not included in the completion or mortality rates.

The mortality rate for the non-MLS participants taken together was 38%. Remarkably, the open call for research participants also produced three participants with self-reported MLS or MLT backgrounds whose data were included with the non-MLS participants because it was not possible to pair the demographic data with the instrument results. Additionally, pairing the demographic data with the results could have enabled participant identification; therefore, the demographic information reported in Table 5 and Table 6 includes data for all students who completed the demographic survey including those participants who opted out of the study since the researcher was unable to remove the anonymous data for those participants.

Table 5

Demographic Information by Age, Gender, Program

	N	Percent
Age		
25+	51	60.71
18-24 years	29	34.52
No response	4	4.76
Gender		
Female	53	63.10
Male	25	29.76
No response	4	4.76
Non-binary / Other	1	1.19
Prefer not to answer	1	1.19
Program		
Medical Laboratory Scientist	50	59.52
Other / No response	32	38.10
Medical Laboratory Technologist	2	2.38

Note. Percentages are based on the total number of participants who started the research study ($N = 84$). Categories are sorted in descending order based on number of responses.

As part of the demographic survey, data were collected to determine participants' experience with online courses and assessments as well as experience with Moodle, the learning management system (LMS) used for this study. Also included in the demographic survey was whether participants had previously been instructed on manual cell counts.

Table 6

Experience with Online Courses, Assessments, and Manual Cell Counts

	N	Percent
Previous an online course		
Yes	77	91.67
No response	4	4.76
No	3	3.57
Previously used Moodle (LMS)		
Yes	56	66.67
No	24	28.57
No response	4	4.76
Previous online exam or assessment		
Yes	79	94.05
No response	4	4.76
No	1	1.19
Previous manual cell count instruction		
No	68	80.95
Yes	12	14.29
No response	4	4.76

Note. Percentages are based on the total number of participants who started the research study (*N* = 84). Categories are sorted in descending order based on number of responses.

Of the participants, approximately 92% had previous experience with an online course and 67% had previously used the same LMS that was used for this study. In addition, 94% of participants reported experience with online exams or assessments. While previous experience with online courses and assessments was an important precursor for this study, participants' prior experience with the content was also important. The majority of participants (81%) did not have experience with manual cell counts, which meant the majority would be seeing the content for the first time, which was ideal for the purposes of this study.

Unfortunately, even though attempts were made to limit any additional instruction regarding manual cell counts in other classes until after the research period, it was discovered that, due to time constraints and other teaching-related factors at the researcher's institution, manual white blood cell (WBC) and red blood cell (RBC) counts were covered in another course on the first day of the first iteration of this research study. This means those study participants were exposed to a live lecture or accessed a recorded lecture prior to starting the treatment intervention. The outside instruction lasted approximately 50 minutes and covered the following components of manual cell counts:

- layout and dimensions of the hemocytometer focused on the grids used for counting white and red blood cells,
- the fact that one millimeter cubed (mm^3) is equal to one microliter (μL),
- how to convert a microliter (μL) to liter (L),
- how to fill (or charge) the hemocytometer chamber,
- how to make a 1 in 20 dilution,
- agreement of cell counts between chambers of 10% agreement (which is not identical to the sperm cell agreement),
- the fact that the 10x microscope objective is used to locate the grid and the cells are counted using the 40x microscope objective,
- the fact that cells that fall on the top and left gridlines are counted and those that fall on the bottom and right gridlines are not counted, and
- the formula for obtaining a cell count.

This additional lecture could be considered a pretraining event (Moreno & Mayer, 2007) because it provided learners with relevant prior knowledge before the online instruction was accessed. The first iteration participants in all four treatment groups accessed the instructional materials for this study after the lecture given by the other MLS instructor. Therefore, this additional instruction could have had an effect on participants' immediate posttest results. Further, participants performed one manual RBC count and one manual WBC count using disposable hemocytometers at one campus location during week seven of the first iteration of the study. Although not all participants in this study performed that lab on that day, 70% of participants completed the delayed post-IMMS measure and delayed knowledge posttest after that lab was held. Having participated in such lab activities prior to completing the delayed knowledge posttest could have had an effect on those results. It is also possible that students reached out to this lab instructor for additional explanation prior to completing the study instruments.

Descriptive Statistics of the Knowledge Instrument

The descriptive statistics for the researcher-developed knowledge instrument are shown in Table 7. The data indicate that the mean scores were similar across the different points in time for all treatment groups and that performance improved for all participants regardless of treatment condition.

Table 7

Descriptive Statistics of Knowledge Scores

	Feedback Type	Scenario Level	Mean	Standard Deviation	*N*
Pretest	Elaborative	No scenario (sE)	3.85	0.77	11
		Scenario (SE)	4.42	0.98	16
		Combined	4.19	0.93	27
	Intrinsic	No scenario (sI)	3.84	0.99	18
		Scenario (SI)	4.04	0.96	18
		Combined	3.94	0.97	36
Immediate Posttest	Elaborative	No scenario (sE)	5.24	1.40	11
		Scenario (SE)	5.41	1.32	16
		Combined	5.34	1.33	27
	Intrinsic	No scenario (sI)	5.29	1.66	18
		Scenario (SI)	5.41	1.58	18
		Combined	5.35	1.60	36
Delayed Posttest	Elaborative	No scenario (sE)	6.13	1.00	11
		Scenario (SE)	5.65	1.32	16
		Combined	5.85	1.20	27
	Intrinsic	No scenario (sI)	5.63	1.85	18
		Scenario (SI)	6.14	1.38	18
		Combined	5.88	1.63	36

Note. Totals represent the removal of three outliers ($N = 63$). The maximum possible score for the knowledge instrument was 10.

As was expected from providing instruction on the topic, the scores for all groups increased between the knowledge pretest and the immediate knowledge posttest;

however, scores also increased for all groups between the immediate posttest and delayed posttest, which was unforeseen. Participants in the scenario with elaborative feedback group (SE) and those in the scenario with intrinsic feedback group (SI) had identical means on the immediate knowledge posttest (*M* = 5.41). As shown in Figure 7, the delayed knowledge posttest means showed a pattern of improved performance for participants who experienced no scenario with elaborative feedback (sE) compared to those who experienced no scenario with intrinsic feedback (sI).

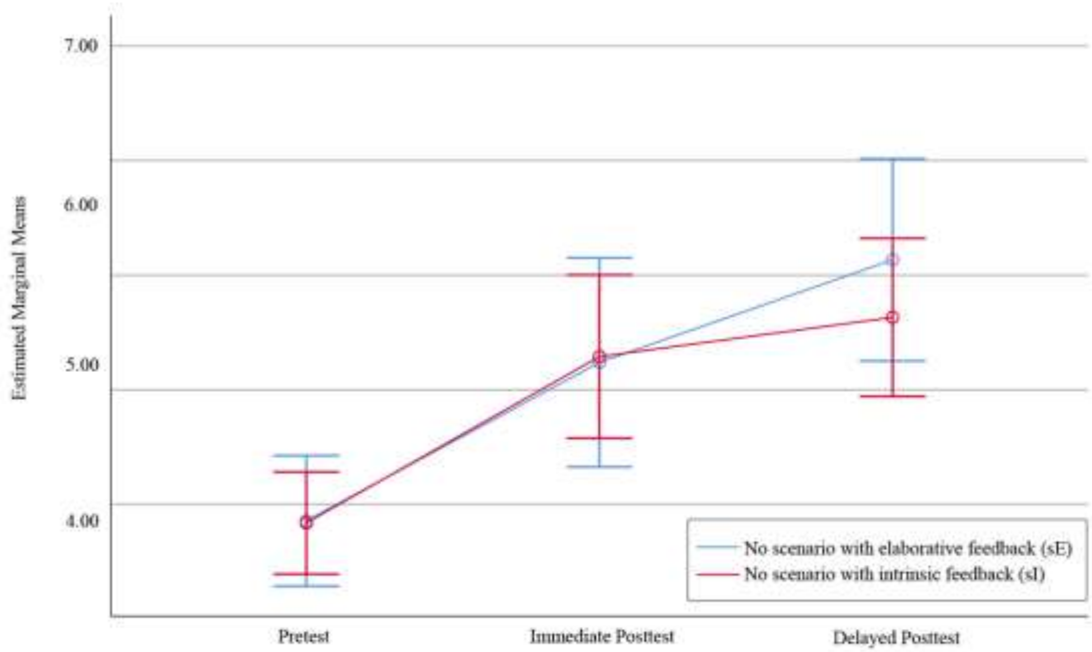


Figure 7. Comparison of knowledge scores by feedback type. Estimated marginal means for participants in the no scenario with elaborative feedback group (sE) and participants in the no scenario with intrinsic feedback group (sI). Error bars represent 95% Confidence Interval.

Further, the participants in the SI group performed better on the delayed knowledge posttest (*M* = 6.14) than those from the SE group (*M* = 5.65) as shown in Figure 8. Additionally, participants in the sE scored similarly to those in the SI group (*M* = 6.13 and *M* = 6.14 respectively) on the delayed knowledge posttest.

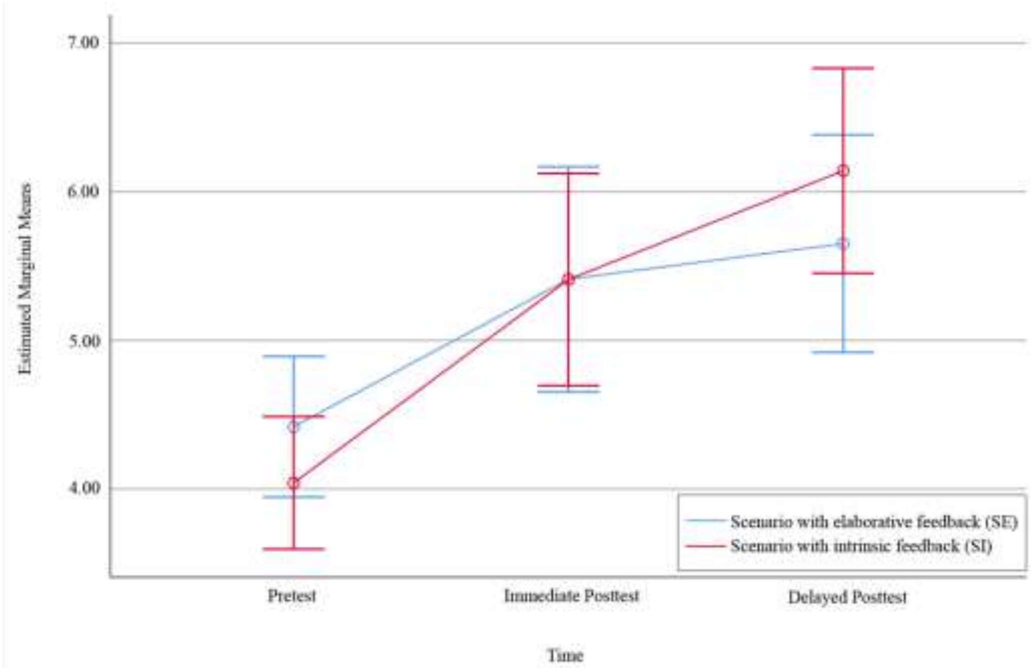


Figure 8. Comparison of knowledge scores by scenario level. Estimated marginal means for participants in the scenario with elaborative feedback group (SE) and participants in the scenario with intrinsic feedback group (SI). Error bars represent 95% Confidence Interval.

Because the relevance of the scenarios and intrinsic feedback used for this study related more to MLS and MLT participants than those in other university programs, the overall knowledge scores were further separated into MLS and non-MLS participants. Participants in the first iteration of the study were known to be students enrolled in the MLS program at the researcher's institution; therefore, the first iteration knowledge scores were separated from those obtained during the second, third, and fourth iterations. The descriptive statistics of the knowledge measure are shown for these sub-groups in Table 8 and Table 9. The mean scores for the MLS participants were similar for all time measurements regardless of feedback type or scenario level. Additionally, the mean scores increased from the knowledge pretest to immediate knowledge posttest and from the immediate knowledge posttest to the delayed knowledge posttest for all MLS participants except for those in the SE group.

Table 8

Descriptive Statistics of MLS Knowledge Scores

	Feedback Type	Scenario Level	Mean	Standard Deviation	*N*
Pretest	Elaborative	No scenario (sE)	4.20	1.14	10
		Scenario (SE)	4.66	1.01	11
		Combined	4.44	1.07	21
	Intrinsic	No scenario (sI)	3.92	0.85	11
		Scenario (SI)	4.38	0.88	11
		Combined	4.15	0.87	22
Immediate Posttest	Elaborative	No scenario (sE)	5.44	1.22	10
		Scenario (SE)	5.81	1.37	11
		Combined	5.63	1.28	21
	Intrinsic	No scenario (sI)	5.57	1.54	11
		Scenario (SI)	5.77	1.76	11
		Combined	5.67	1.61	22
Delayed Posttest	Elaborative	No scenario (sE)	6.19	1.06	10
		Scenario (SE)	5.79	1.38	11
		Combined	5.98	1.23	21
	Intrinsic	No scenario (sI)	6.05	1.13	11
		Scenario (SI)	6.46	1.62	11
		Combined	6.26	1.38	22

Note. Totals represent the removal of two outliers (N = 43).

Similar to the data from all participants (see Table 7), the delayed knowledge posttest results show that MLS participants in the sE group scored higher (M = 6.19) than MLS participants in the SE group (M = 5.79). Similarly, MLS participants in the SI group scored higher (M = 6.46) on the delayed knowledge posttest than MLS participants in the sI group (M = 6.05). MLS participants in the SE group scored the highest on the immediate knowledge posttest (M = 5.81), whereas MLS participants in the SI group scored highest on the delayed knowledge posttest (M = 6.46).

The descriptive statistics in Table 9 show that mean scores were similar for all non-MLS participants for the three time measurements. Additionally, the data show that non-MLS participants improved from the knowledge pretest to the immediate knowledge posttest and also from the immediate knowledge posttest to the delayed knowledge posttest. Non-MLS participants in the sE group scored higher (M = 4.75) than non-MLS participants in the SE group (M = 4.54) on the immediate knowledge posttest as well as on the delayed knowledge posttest (M = 6.48 and M = 5.34 respectively). Similarly, non-MLS participants in the SI group scored higher (M = 4.84) than non-MLS participants in the sI group (M = 4.79) on the immediate knowledge posttest. This pattern also continued to the delayed knowledge posttest with a mean score of 5.64 for non-MLS participants in the SI group and 5.46 for non-MLS participants in the sI group.

Table 9

Descriptive Statistics of Non-MLS Knowledge Scores

	Feedback Type	Scenario Level	Mean	Standard Deviation	*N*
Pretest	Elaborative	No scenario (sE)	3.48	0.06	2
		Scenario (SE)	3.89	0.72	5
		Combined	3.77	0.62	7
	Intrinsic	No scenario (sI)	3.74	1.36	6
		Scenario (SI)	3.50	0.88	7
		Combined	3.61	1.08	13
Immediate Posttest	Elaborative	No scenario (sE)	4.75	2.50	2
		Scenario (SE)	4.54	0.66	5
		Combined	4.60	1.16	7
	Intrinsic	No scenario (sI)	4.79	2.05	6
		Scenario (SI)	4.84	1.13	7
		Combined	4.81	1.55	13
Delayed Posttest	Elaborative	No scenario (sE)	6.48	1.10	2
		Scenario (SE)	5.34	1.24	5
		Combined	5.67	1.24	7
	Intrinsic	No scenario (sI)	5.46	2.43	6
		Scenario (SI)	5.64	0.72	7
		Combined	5.56	1.66	13

Note. Totals represent the removal of one outlier (N = 20).

It is curious that the pattern of increased scores between the immediate and delayed posttest scores that was seen for the MLS participants was also observed for the non-MLS participants. Because the majority (83%) of participants in the second, third, and fourth iterations reported they were not enrolled in MLS or MLT programs, it was unlikely that additional instruction was the reason for their increased scores as could be the case for the MLS participants from the researcher's university. Among MLS participants, the SI group scored highest on the delayed knowledge posttest (*M* = 6.46). However, among non-MLS participants, the sE group scored the highest on the delayed knowledge posttest (*M* = 6.48).

Descriptive Statistics of the IMMS Instrument

The descriptive statistics of Keller's (2010) IMMS scores for all participants are shown in Table 10. The data indicate that the mean scores were similar for all participants across all time measurements regardless of treatment condition. The data also show that IMMS ratings decreased from the pre-IMMS measurement to the immediate post-IMMS measurement, but increased between the immediate and delayed post-IMMS measurements for all participants except those in the SE group.

Table 10

Descriptive Statistics of IMMS Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pre-IMMS	No scenario	Elaborative (sE)	143.23	21.54	13
		Intrinsic (sI)	140.76	18.94	17
		Combined	141.83	19.79	30
	Scenario	Elaborative (SE)	140.73	16.44	15
		Intrinsic (SI)	138.95	16.95	19
		Combined	139.74	16.50	34
Immediate Post-IMMS	No scenario	Elaborative (sE)	140.23	22.96	13
		Intrinsic (sI)	135.35	22.12	17
		Combined	137.47	22.23	30
	Scenario	Elaborative (SE)	131.53	24.85	15
		Intrinsic (SI)	129.58	29.48	19
		Combined	130.44	27.15	34
Delayed Post-IMMS	No scenario	Elaborative (sE)	142.00	20.04	13
		Intrinsic (sI)	139.76	22.96	17
		Combined	140.73	21.41	30
	Scenario	Elaborative (SE)	131.20	22.25	15
		Intrinsic (SI)	135.58	23.20	19
		Combined	133.65	22.55	34

Note. Totals represent the removal of two outliers ($N = 64$). The minimum and maximum possible scores were 36 and 180 respectively.

Prior to receiving any instruction or feedback, participants in the sE group rated somewhat higher on the pre-IMMS ($M = 143.23$) than their counterparts in the sI group ($M = 140.76$). Similarly, the participants in the SE group had higher ratings ($M = 140.73$) than the participants in the SI group ($M = 138.95$). The SE and sI groups had nearly identical mean scores on the pre-IMMS ($M = 140.73$ and $M = 140.76$ respectively). Participants in the sE group had the highest ratings on all of IMMS measurements.

To examine whether the relevance of the instruction had more of an effect for MLS participants, the data were split into MLS and non-MLS participants for analysis. The descriptive statistics for MLS participants presented in Table 11 indicate that the mean scores were similar for all participants across all time measurements regardless of treatment condition with the exception of the nearly ten-point difference in scores on the delayed post-IMMS measurement between the SE (*M* = 130.45) and SI groups (*M* = 140.00). The data also show that IMMS scores decreased from the pre-IMMS measurement to the immediate post-IMMS measurement, but then increased slightly between the immediate and delayed post-IMMS measurements for MLS participants in the sI and SI groups, but decreased again for MLS participants in the sE and SE groups. Notably, the no scenario participants (sE and sI) reported higher motivation levels on the pre-IMMS (*M*= 145.80 and *M* = 150.18 respectively) than those in the SE group (*M* = 138.18) and SI group (*M* = 142.00) prior to experiencing the treatments.

Table 11

Descriptive Statistics of MLS IMMS Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pre-IMMS	No scenario	Elaborative (sE)	145.80	19.75	10
		Intrinsic (sI)	150.18	12.31	11
		Combined	148.10	16.01	21
	Scenario	Elaborative (SE)	138.18	17.84	11
		Intrinsic (SI)	142.00	10.56	9
		Combined	139.90	14.78	20
Immediate Post-IMMS	No scenario	Elaborative (sE)	145.50	10.93	10
		Intrinsic (sI)	141.64	22.25	11
		Combined	143.48	17.47	21
	Scenario	Elaborative (SE)	132.27	28.48	11
		Intrinsic (SI)	125.44	34.15	9
		Combined	129.20	30.50	20
Delayed Post-IMMS	No scenario	Elaborative (sE)	144.80	16.70	10
		Intrinsic (sI)	146.09	22.27	11
		Combined	145.48	19.34	21
	Scenario	Elaborative (SE)	130.45	24.82	11
		Intrinsic (SI)	140.00	17.30	9
		Combined	134.75	21.77	20

Note. Totals represent the removal of four outliers (*N* = 41). The minimum and maximum possible scores were 36 and 180 respectively.

The descriptive statistics for non-MLS participants presented in Table 12 indicate that participants in the sE group scored higher (*M* = 150.50) than those in the sI group (*M* = 126.80) even though participants had not yet been introduced to the treatment conditions.

Table 12

Descriptive Statistics of Non-MLS IMMS Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pre-IMMS	No scenario	Elaborative (sE)	150.50	16.26	2
		Intrinsic (sI)	126.80	16.86	5
		Combined	133.57	19.16	7
	Scenario	Elaborative (SE)	147.75	10.50	4
		Intrinsic (SI)	129.43	10.41	7
		Combined	136.09	13.54	11
Immediate Post-IMMS	No scenario	Elaborative (sE)	145.00	32.53	2
		Intrinsic (sI)	127.20	18.06	5
		Combined	132.28	21.66	7
	Scenario	Elaborative (SE)	129.50	13.02	4
		Intrinsic (SI)	124.28	24.84	7
		Combined	126.18	20.69	11
Delayed Post-IMMS	No scenario	Elaborative (sE)	145.00	32.53	2
		Intrinsic (sI)	134.60	15.66	5
		Combined	137.57	19.12	7
	Scenario	Elaborative (SE)	133.25	15.78	4
		Intrinsic (SI)	122.28	23.59	7
		Combined	126.27	20.96	11

Note. Totals represent the removal of three outliers (*N* = 18). The minimum and maximum possible scores were 36 and 180 respectively.

Similarly, the SE participants scored higher on the pre-IMMS measure (*M* = 147.75) than the participants in the SI group (*M* = 129.43) prior to experiencing the treatments. Additionally, there was a difference of almost 18 points between the immediate post-IMMS scores of participants in the sE group (*M* = 145.00) and those in the sI group (*M* = 127.20). Further, participants in the sE group had no change in mean score between the immediate post-IMMS and delayed post-IMMS measures.

Descriptive statistics of the IMMS subscales. Keller's (2010) IMMS instrument includes questions related to the attention, relevance, confidence, and satisfaction (ARCS) subscales. Due to the lack of statistically significant results, the subscales were examined to determine if a change in one subscale was obscured by an opposite change in a different subscale. Table 13 shows the mean scores for all study participants on the attention subscale. According to Keller, the attention category of the ARCS model refers to capturing and sustaining the learner's attention; therefore, the items on the attention subscale relate to "boredom and lack of stimulation" (p. 282) with instructional writing and design.

Table 13

Descriptive Statistics of Attention Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	48.31	7.89	13
		Intrinsic (sI)	47.50	7.57	18
		Combined	47.84	7.58	31
	Scenario	Elaborative (SE)	48.53	7.03	15
		Intrinsic (SI)	48.00	5.50	19
		Combined	48.24	6.12	34
Immediate Posttest Score	No scenario	Elaborative (sE)	47.31	6.80	13
		Intrinsic (sI)	43.94	9.98	18
		Combined	45.35	8.82	31
	Scenario	Elaborative (SE)	43.67	9.75	15
		Intrinsic (SI)	42.47	9.70	19
		Combined	43.00	9.59	34
Delayed Posttest Score	No scenario	Elaborative (sE)	46.85	8.11	13
		Intrinsic (sI)	44.67	9.36	18
		Combined	45.58	8.78	31
	Scenario	Elaborative (SE)	44.20	8.54	15
		Intrinsic (SI)	44.74	7.94	19
		Combined	44.50	8.09	34

Note. Totals represent the removal of one outlier ($N = 65$). The minimum and maximum possible scores were 12 and 60 respectively.

The data indicate that participant ratings in all groups decreased from the pre-IMMS to the immediate post-IMMS and all participant ratings increased slightly between the immediate post-IMMS and delayed post-IMMS except for participants in the sE group; however, the sE group had the highest attention ratings on both the immediate and delayed IMMS surveys (*M* = 47.31 and *M* = 46.85 respectively).

Table 14 shows the data for all participants on the relevance subscale of the IMMS instrument. Keller's (2010) relevance category pertains to perceived usefulness of the instruction. The more attainable and useful a learner finds the instruction, the higher their commitment to accomplish it; therefore, the relevance subscale is intended to measure the learners' perceived personal connection to the instruction.

Table 14

Descriptive Statistics of Relevance Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	37.33	4.64	12
		Intrinsic (sI)	36.06	5.46	16
		Combined	36.07	5.07	28
	Scenario	Elaborative (SE)	34.06	5.46	16
		Intrinsic (SI)	34.21	5.20	19
		Combined	34.14	5.24	35
Immediate Posttest Score	No scenario	Elaborative (sE)	36.00	5.09	12
		Intrinsic (sI)	35.50	5.08	16
		Combined	35.71	5.00	28
	Scenario	Elaborative (SE)	31.69	7.59	16
		Intrinsic (SI)	34.26	6.52	19
		Combined	33.08	7.04	35
Delayed Posttest Score	No scenario	Elaborative (sE)	36.92	4.29	12
		Intrinsic (sI)	35.94	5.47	16
		Combined	36.35	4.94	28
	Scenario	Elaborative (SE)	31.75	8.00	16
		Intrinsic (SI)	33.63	6.40	19
		Combined	32.77	7.13	35

Note. Totals represent the removal of three outliers (*N* = 63). The minimum and maximum possible scores were nine and 45 respectively.

The results in Table 14 show that the mean relevance scores for all participants decreased between the pre-IMMS and the immediate post-IMMS measures except for participants in the SI group where the mean increased slightly; however, the SI participants' mean scores were the only such scores to decrease slightly between the immediate post-IMMS and delayed post-IMMS measures.

The group mean scores for all participants on the confidence subscale are shown in Table 15. According to Keller (2010), the confidence category of the ARCS model relates to learners' beliefs that they can succeed. The confidence subscale on the IMMS measures learners' perception of their likelihood of success with the instruction.

Table 15

Descriptive Statistics of Confidence Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	34.92	5.02	13
		Intrinsic (sI)	35.00	3.97	17
		Combined	34.97	4.37	30
	Scenario	Elaborative (SE)	34.80	4.74	15
		Intrinsic (SI)	33.26	5.54	19
		Combined	33.94	5.18	34
Immediate Posttest Score	No scenario	Elaborative (sE)	36.54	6.09	13
		Intrinsic (sI)	34.53	6.52	17
		Combined	35.40	6.31	30
	Scenario	Elaborative (SE)	35.47	5.93	15
		Intrinsic (SI)	31.68	9.08	19
		Combined	33.35	7.97	34
Delayed Posttest Score	No scenario	Elaborative (sE)	36.23	5.40	13
		Intrinsic (sI)	36.41	5.28	17
		Combined	36.33	5.24	30
	Scenario	Elaborative (SE)	34.47	5.91	15
		Intrinsic (SI)	34.74	5.44	19
		Combined	34.62	5.57	34

Note. Totals represent the removal of two outliers (N = 64). The minimum and maximum possible scores were nine and 45 respectively.

The confidence subscale results indicate that participants who received elaborative feedback (sE and SE) had increased confidence from the pre-IMMS to the immediate post-IMMS, but decreased confidence between the immediate post-IMMS and the delayed post-IMMS. In contrast, participants who received intrinsic feedback (sI and SI) had decreased ratings from the pre-IMMS to the immediate post-IMMS, but reported increased confidence between the immediate post-IMMS and the delayed post-IMMS.

Table 16 shows the group means on the satisfaction subscale for all participants. Keller (2010) stated, “intrinsic satisfaction, can result from feelings of mastery and from the pleasure of having succeeded at a task which was meaningful and challenging” (p. 166); therefore the satisfaction subscale is a measure of learners’ feelings and attitudes toward their performance and their expectations.

Table 16

Descriptive Statistics of Satisfaction Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	23.38	4.98	13
		Intrinsic (sI)	22.44	5.19	18
		Combined	22.84	5.04	31
	Scenario	Elaborative (SE)	22.56	4.32	16
		Intrinsic (SI)	23.47	4.26	19
		Combined	23.05	4.24	35
Immediate Posttest Score	No scenario	Elaborative (sE)	22.15	6.36	13
		Intrinsic (sI)	20.33	6.78	18
		Combined	21.10	6.56	31
	Scenario	Elaborative (SE)	19.00	6.69	16
		Intrinsic (SI)	21.16	8.02	19
		Combined	20.17	7.41	35
Delayed Posttest Score	No scenario	Elaborative (sE)	22.62	4.72	13
		Intrinsic (sI)	22.22	6.00	18
		Combined	22.39	5.42	31
	Scenario	Elaborative (SE)	19.12	6.26	16
		Intrinsic (SI)	22.47	6.38	19
		Combined	20.94	6.46	35

Note. There were no outliers ($N = 66$). The minimum and maximum possible scores were six and 30 respectively.

Table 16 shows that all participants, regardless of treatment group, reported decreased satisfaction from the pre-IMMS to the immediate post-IMMS. Conversely, all participants, regardless of treatment group rated their satisfaction slightly higher between the immediate post-IMMS and delayed post-IMMS.

Descriptive statistics of MLS IMMS subscales. The overall IMMS scores and subscale scores were further divided and analyzed for patterns among the MLS and non-MLS participants. The subscale results for the MLS participants will be discussed next, followed by a discussion of the subscale results for the non-MLS participants. Table 17 shows the mean scores for MLS participants on the attention subscale of the IMMS.

Table 17

Descriptive Statistics of MLS Attention Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	47.91	8.07	11
		Intrinsic (sI)	49.25	7.72	12
		Combined	48.61	7.74	23
	Scenario	Elaborative (SE)	46.54	7.17	11
		Intrinsic (SI)	48.09	6.04	11
		Combined	47.32	6.52	22
Immediate Posttest Score	No scenario	Elaborative (sE)	46.27	6.54	11
		Intrinsic (sI)	45.67	10.22	12
		Combined	45.96	8.47	23
	Scenario	Elaborative (SE)	41.72	10.61	11
		Intrinsic (SI)	42.27	10.78	11
		Combined	42.00	10.44	22
Delayed Posttest Score	No scenario	Elaborative (sE)	46.09	8.07	11
		Intrinsic (sI)	46.08	10.32	12
		Combined	46.08	9.10	23
	Scenario	Elaborative (SE)	42.82	9.70	11
		Intrinsic (SI)	46.18	7.30	11
		Combined	44.50	8.56	22

Note. There were no outliers (*N* = 45). The minimum and maximum possible scores were 12 and 60 respectively.

The data in Table 17 show that MLS participants in all treatment groups reported lower attention scores on the immediate post-IMMS than they did on the pre-IMMS; however, the MLS participants rated their attention slightly higher on the delayed post-IMMS than the immediate post-IMMS except for MLS participants in the sE group; however, MLS participants in the sE group had the highest attention ratings on the immediate post-IMMS (*M* = 46.27). In contrast, the MLS participants in the SI group had the highest attention ratings on the delayed post-IMMS (*M* = 46.18).

Table 18

Descriptive Statistics of MLS Relevance Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	37.40	5.12	10
		Intrinsic (sI)	37.92	3.92	12
		Combined	37.68	4.40	22
	Scenario	Elaborative (SE)	35.45	5.37	11
		Intrinsic (SI)	35.80	3.97	10
		Combined	35.62	4.64	21
Immediate Posttest Score	No scenario	Elaborative (sE)	36.50	5.23	10
		Intrinsic (sI)	36.83	4.78	12
		Combined	36.68	4.87	22
	Scenario	Elaborative (SE)	34.18	7.14	11
		Intrinsic (SI)	37.10	4.56	10
		Combined	35.57	6.09	21
Delayed Posttest Score	No scenario	Elaborative (sE)	37.00	4.69	10
		Intrinsic (sI)	36.33	6.05	12
		Combined	36.64	5.36	22
	Scenario	Elaborative (SE)	33.73	8.27	11
		Intrinsic (SI)	35.70	5.01	10
		Combined	34.67	6.82	21

Note. Totals represent the removal of two outliers (N = 43). The minimum and maximum possible scores were nine and 45 respectively.

On the relevance subscale, Table 18 shows that all MLS participants except those in the SI group had decreased ratings between the pre-IMMS and immediate post-IMMS. All MLS participant ratings except for those in the sE group decreased between the immediate post-IMMS and delayed post-IMMS. MLS participants in the SI group had the highest immediate post-IMMS ratings on the relevance subscale (M = 37.10), whereas, the MLS participants in the sE group had the highest relevance subscale ratings on the delayed post-IMMS measure (M = 37.00).

Table 19 shows the mean scores for MLS participants on the confidence subscale of the IMMS. The results show that MLS participants in the SI group had the highest confidence ratings on all three IMMS measures.

Table 19

Descriptive Statistics of MLS Confidence Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	33.40	4.22	10
		Intrinsic (sI)	34.92	3.78	12
		Combined	34.23	3.96	22
	Scenario	Elaborative (SE)	33.45	4.06	11
		Intrinsic (SI)	35.43	5.44	7
		Combined	34.22	4.60	18
Immediate Posttest Score	No scenario	Elaborative (sE)	36.60	5.93	10
		Intrinsic (sI)	35.08	5.71	12
		Combined	35.77	5.72	22
	Scenario	Elaborative (SE)	36.27	5.48	11
		Intrinsic (SI)	38.58	3.95	7
		Combined	37.17	4.95	18
Delayed Posttest Score	No scenario	Elaborative (sE)	35.90	4.79	10
		Intrinsic (sI)	36.25	5.14	12
		Combined	36.09	4.87	22
	Scenario	Elaborative (SE)	34.54	4.48	11
		Intrinsic (SI)	38.14	4.52	7
		Combined	35.94	4.72	18

Note. Totals represent the removal of five outliers (N = 40). The minimum and maximum possible scores were nine and 45 respectively.

As shown in Table 19, MLS participants' confidence ratings increased from the pre-IMMS to the immediate post-IMMS regardless of treatment group. The only MLS participants with increased confidence ratings between the immediate post-IMMS and delayed post-IMMS were those in the sI group.

The results for the MLS participants' satisfaction subscale of the IMMS are presented in Table 20. As was the case for all participants, the group means for MLS participants on the satisfaction subscale decreased from pre-IMMS to immediate post-IMMS.

Table 20

Descriptive Statistics of MLS Satisfaction Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	23.80	4.61	10
		Intrinsic (sI)	24.50	4.12	12
		Combined	24.18	4.26	22
	Scenario	Elaborative (SE)	22.73	4.82	11
		Intrinsic (SI)	24.45	4.52	11
		Combined	23.59	4.65	22
Immediate Posttest Score	No scenario	Elaborative (sE)	23.70	3.59	10
		Intrinsic (sI)	21.17	6.73	12
		Combined	22.32	5.56	22
	Scenario	Elaborative (SE)	20.09	6.85	11
		Intrinsic (SI)	20.27	9.31	11
		Combined	20.18	7.97	22
Delayed Posttest Score	No scenario	Elaborative (sE)	23.10	4.36	10
		Intrinsic (sI)	23.25	5.48	12
		Combined	23.18	4.88	22
	Scenario	Elaborative (SE)	19.36	6.23	11
		Intrinsic (SI)	22.82	6.75	11
		Combined	21.09	6.58	22

Note. Totals represent the removal of one outlier (N = 44). The minimum and maximum possible scores were six and 30 respectively.

MLS participants who received elaborative feedback (sE and SE) had lower satisfaction scores from the immediate post-IMMS to the delayed post-IMMS; however, MLS participants who received intrinsic feedback (sI and SI) rated their satisfaction slightly higher from the immediate post-IMMS to the delayed post-IMMS. MLS participants in the sE group had the highest satisfaction ratings on the immediate post-IMMS while MLS participants in the sI group had the highest satisfaction ratings on the pre-IMMS and delayed post-IMMS measures.

Descriptive statistics of non-MLS IMMS subscales. Table 21 shows non-MLS participants’ mean scores for the IMMS attention subscale. Non-MLS participants in the SE groups reported the highest attention scores for all three IMMS measurements.

Table 21

Descriptive Statistics of non-MLS Attention Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	44.00	-	1
		Intrinsic (sI)	44.00	6.45	6
		Combined	44.00	5.89	7
	Scenario	Elaborative (SE)	54.00	2.16	4
		Intrinsic (SI)	46.28	2.50	7
		Combined	49.09	4.50	11
Immediate Posttest Score	No scenario	Elaborative (sE)	48.00	-	1
		Intrinsic (sI)	40.50	9.35	6
		Combined	41.57	9.00	7
	Scenario	Elaborative (SE)	49.00	4.08	4
		Intrinsic (SI)	43.00	9.38	7
		Combined	45.18	8.18	11
Delayed Posttest Score	No scenario	Elaborative (sE)	44.00	-	1
		Intrinsic (sI)	41.83	7.03	6
		Combined	42.14	6.47	7
	Scenario	Elaborative (SE)	48.00	0.82	4
		Intrinsic (SI)	41.43	8.64	7
		Combined	43.82	7.48	11

Note. Totals represent the removal of three outliers (N = 18). The minimum and maximum possible scores were 36 and 180 respectively.

The data in Table 21 show that all non-MLS participants' attention ratings decreased from the pre-IMMS to the immediate post-IMMS except for those participants in the sE group. Additionally, all scores decreased from the immediate post-IMMS to the delayed post-IMMS for non-MLS participants except for those in the sI group.

Non-MLS participants' scores on the IMMS relevance subscale are reported in Table 22. Non-MLS participants in the sE group reported the highest relevance ratings on all three IMMS measures.

Table 22

Descriptive Statistics of non-MLS Relevance Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	37.00	0.00	2
		Intrinsic (sI)	28.67	6.09	6
		Combined	30.75	6.43	8
	Scenario	Elaborative (SE)	32.00	4.83	4
		Intrinsic (SI)	30.28	4.61	7
		Combined	30.91	4.53	11
Immediate Posttest Score	No scenario	Elaborative (sE)	35.50	4.95	2
		Intrinsic (sI)	27.17	7.96	6
		Combined	28.75	7.57	8
	Scenario	Elaborative (SE)	28.25	4.11	4
		Intrinsic (SI)	30.43	5.97	7
		Combined	29.64	5.26	11
Delayed Posttest Score	No scenario	Elaborative (sE)	36.50	2.12	2
		Intrinsic (sI)	29.83	8.13	6
		Combined	31.50	7.58	8
	Scenario	Elaborative (SE)	29.75	2.99	4
		Intrinsic (SI)	28.71	5.50	7
		Combined	29.09	4.59	11

Note. Totals represent the removal of two outliers (*N* = 19). The minimum and maximum possible scores were nine and 45 respectively.

Non-MLS participants in all treatment groups except the SI group reported lower relevance ratings on the immediate post-IMMS than the pre-IMMS. On the contrary, non-MLS participants in the SI group were the only participants who did not report increased ratings from the immediate post-IMMS to the delayed post-IMMS.

Table 23 shows the IMMS confidence subscale results for the non-MLS participants. As for the relevance subscale with non-MLS participants, the highest confidence ratings on all three IMMS measures were from non-MLS participants in the sE group.

Table 23

Descriptive Statistics of non-MLS Confidence Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	37.50	2.12	2
		Intrinsic (sI)	34.00	4.69	4
		Combined	35.17	4.17	6
	Scenario	Elaborative (SE)	36.33	3.06	3
		Intrinsic (SI)	32.71	5.34	7
		Combined	33.80	4.92	10
Immediate Posttest Score	No scenario	Elaborative (sE)	36.00	11.31	2
		Intrinsic (sI)	31.00	8.41	4
		Combined	32.67	8.64	6
	Scenario	Elaborative (SE)	29.67	2.31	3
		Intrinsic (SI)	31.57	5.83	7
		Combined	31.00	4.97	10
Delayed Posttest Score	No scenario	Elaborative (sE)	36.00	11.31	2
		Intrinsic (sI)	35.00	5.48	4
		Combined	35.33	6.62	6
	Scenario	Elaborative (SE)	31.00	9.00	3
		Intrinsic (SI)	33.86	4.30	7
		Combined	33.00	5.68	10

Note. Totals represent the removal of five outliers (N = 16). The minimum and maximum possible scores were nine and 45 respectively.

The means in Table 23 indicate that all non-MLS participants' confidence ratings decreased from the pre-IMMS to the immediate post-IMMS. With the exception of non-MLS participants in the sE group, the non-MLS participants' confidence ratings increased slightly from the immediate post-IMMS to the delayed post-IMMS measure.

The IMMS satisfaction subscale results for non-MLS participants are shown in Table 24. The satisfaction ratings for non-MLS participants who received elaborative feedback (sE and SE) decreased between the pre-IMMS and immediate post-IMMS measures.

Table 24

Descriptive Statistics of non-MLS Satisfaction Subscale Scores

	Scenario Level	Feedback Type	Mean	Standard Deviation	*N*
Pretest Score	No scenario	Elaborative (sE)	25.50	4.95	2
		Intrinsic (sI)	18.33	4.88	6
		Combined	20.12	5.62	8
	Scenario	Elaborative (SE)	22.20	3.42	5
		Intrinsic (SI)	22.12	3.72	8
		Combined	22.15	3.46	13
Immediate Posttest Score	No scenario	Elaborative (sE)	22.50	9.19	2
		Intrinsic (sI)	18.67	7.20	6
		Combined	19.62	7.23	8
	Scenario	Elaborative (SE)	16.60	6.35	5
		Intrinsic (SI)	22.38	6.21	8
		Combined	20.15	6.67	13
Delayed Posttest Score	No scenario	Elaborative (sE)	21.50	9.19	2
		Intrinsic (sI)	20.17	7.00	6
		Combined	20.50	6.89	8
	Scenario	Elaborative (SE)	18.60	7.02	5
		Intrinsic (SI)	22.00	6.26	8
		Combined	20.69	6.50	13

Note. There were no outliers (N = 21). The minimum and maximum possible scores were six and 30 respectively.

Similarly, non-MLS participants who received intrinsic feedback (sI and SI) reported higher satisfaction scores on the immediate post-IMMS measure than on the pre-IMMS measure. On the other hand, non-MLS participants in the sI and SE groups reported higher satisfaction while those in the sE and SI groups reported lower satisfaction on the delayed post-IMMS measure than on the immediate post-IMMS measure.

Inferential Statistics

Research Questions One and Two were examined using a pair of 2×2 repeated measures ANOVAs with between group factors and random assignment. Additionally, as part of research question one, a chi-square test of independence was used to determine whether participants' mastery of manually counting sperm cells was related to the treatment conditions. Participants who answered at least two out of three manual cell counts correctly were considered to have mastered the ability. On the other hand, participants who answered one or fewer manual cell counts correctly were considered not to have mastered the skill. To qualify as correct, participants' cell counts were required to be within ±1 standard deviation (SD) of the mean of expert counts. The results for each research question are presented next.

Research Question One. Research Question One asked whether the use of SBeL and feedback type have a significant effect over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, posttest, and delayed performance posttest. Research question one and its three sub-parts are included below with the corresponding null and alternate hypotheses:

1. Does the use of SBeL and feedback type have a significant effect over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest?
    a. Does the use of SBeL have a significant main effect over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest? The null hypothesis was that there would not be a difference in performance based on scenario level (H_0: $\mu_{\text{Scenario}} = \mu_{\text{No scenario}}$) and the alternate hypothesis

was that a difference would exist between participants' performance on the knowledge measurement based on scenario level (H_1: $\mu_{\text{Scenario}} \neq \mu_{\text{No scenario}}$).

b. Does feedback type have a significant main effect over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest? The null hypothesis was that there would not be a difference in performance based on feedback type (H_0: $\mu_{\text{Elaborative}} = \mu_{\text{Intrinsic}}$) and the alternate hypothesis was that a difference would exist between participants' performance on the knowledge measurement over time based on feedback type (H_1: $\mu_{\text{Elaborative}} \neq \mu_{\text{Intrinsic}}$).

c. Is there a significant interaction effect of the combination of SBeL and feedback type over time on college students' ability to manually count sperm cells as measured by the comparison of participant results on the researcher-developed online pretest, immediate posttest, and delayed performance posttest? The null hypothesis was that there would not be an interaction between scenario level and feedback type and the main effects would explain all of the mean differences. On the other hand, the alternate hypothesis was that there would be an interaction between the two independent variables.

Certain conditions and assumptions must be met in order to make accurate inferences from the 2×2 repeated measures ANOVA (Gravetter & Wallnau, 2007). First, the repeated measures ANOVA requires a continuous dependent variable, one categorical within-subjects factor with at least two levels, and at least one categorical between-subjects factor with two or more levels. The dependent variable for research question one was scores on the researcher-developed knowledge instrument, which was a continuous variable. The study also included the within-subjects factor of time with three levels (pre-, immediate post-, and delayed posttest). Lastly, the study included two categorical between-subjects factors (scenario level and feedback type), each with two levels (present/not present and elaborative/intrinsic respectively).

After meeting the three initial conditions, the data were analyzed for outliers, normal distribution of the dependent variable, and homogeneity of variances and covariances, and sphericity. There were three outliers as assessed by boxplot, which were removed prior to analysis. No additional outliers were identified by boxplot after removal of the initial outliers. The knowledge scores were normally distributed as assessed by Shapiro-Wilk's test of normality on the studentized residuals ($p > .05$). Levene's test of homogeneity of variances and Box's M test showed there was homogeneity of variances ($p > .05$) and covariances ($p > .05$) respectively. The assumption of sphericity was also met as indicated by Mauchly's test of sphericity, $\chi^2(2) = 0.45$, $p = .800$. The results of the repeated measures ANOVA indicated there was not a statistically significant effect of feedback type over time on participants' knowledge test scores, $F(2, 118) = 0.22$, $p = .800$, partial $\eta^2 = .004$. There was also no statistically significant effect of scenario level over time on participants' knowledge test scores, $F(2, 118) = 0.55$, $p = .581$, partial $\eta^2 = .009$. There was no statistically significant interaction between scenario level and feedback type over time on performance, $F(2, 118) = 1.94$, $p = .148$, partial $\eta^2 = .032$.

The main effect of time did show a statistically significant difference in participants' knowledge test performance at the three different time points, $F(2, 118) = 55.86$, $p < .001$, partial $\eta^2 = .486$. There was a statistically significant increase in knowledge test scores from the pretest ($M = 4.04$) to the immediate posttest ($M = 5.34$, 95% CI, 0.84 to 1.76, $p < .001$). Additionally, there was a statistically significant increase in knowledge test scores from the immediate posttest ($M = 5.34$) to the delayed posttest ($M = 5.89$, 95% CI, 0.12 to 0.99, $p < .001$). There was also a significant increase from the pretest to the delayed posttest (95% CI, 1.42 to 2.28, $p < .001$).

Because a small sample size was obtained, lack of power was a concern; therefore, two 2×2 factorial ANOVAs were used to examine the differences between scenario levels and feedback types on the immediate and delayed posttests separately to check that power was not lost as part of the repeated measures ANOVA. There are six requirements and assumptions for a 2×2 factorial ANOVA (Gravetter & Wallnau, 2007). The first two requirements are a continuous dependent variable and two independent variables with two categorical levels. This study examined continuous knowledge scores for two independent variables (scenario level and feedback type) with two levels each. The third assumption requires independence of observations, which was met for this study because participants only participated in one treatment group during each posttest. To meet the fourth assumption, there should not be any significant outliers in the data. The fifth and sixth assumptions require the residuals of the dependent variable to be normally distributed and have equal variances. There were no outliers for the immediate posttest scores, as assessed by inspection of the boxplots and the data were normally distributed as indicated by Shapiro-Wilk's test ($p > .05$). Further, Levene's test for equality of variances showed there was homogeneity of variances ($p > .05$).

An examination of the knowledge pretest scores indicated there were no statistically significant differences of feedback type $F(1, 62) = 0.49$, $p = .485$, partial $\eta^2 = .008$, power = 0.11 or scenario level $F(1, 62) = 2.30$, $p = .134$, partial $\eta^2 = .036$, power = 0.34. Similarly, there was not a statistically significant interaction between feedback and scenario on the knowledge pretest scores $F(1, 62) = 0.53$, $p = .469$, partial $\eta^2 = .008$, power = 0.11.

There was not a statistically significant interaction between feedback type and scenario level on the immediate posttest scores, $F(1, 62) = 0.06$, $p = .815$, partial $\eta^2 = .001$, power = 0.06. The main effect of feedback type on mean immediate posttest scores was also not statistically significant, $F(1, 62) = 0.01$, $p = .942$, partial $\eta^2 = .000$, power = 0.05. Similarly, the main effect of scenario level on mean immediate posttest scores was not statistically significant, $F(1, 62) = 0.16$, $p = .689$, partial $\eta^2 = .003$, power = 0.07.

An examination of the boxplots for the delayed posttest residuals revealed one outlier for the SI group; however, the data point was not removed for analysis because removing outliers may create additional outliers and continuing to remove data points could bias the results (T. Peterson, personal communication, June 1, 2020). The Shapiro-Wilk's test indicated the data were normally distributed ($p > .05$) and Levene's test for equality of variances showed there was homogeneity of variances ($p > .05$).

There was not a statistically significant interaction between feedback type and scenario level for the delayed posttest scores, $F(1, 62) = 0.62$, $p = .434$, partial $\eta^2 = .010$, power = 0.13. Additionally, there were no statistically significant main effects for scenario level or feedback type on the delayed posttest, $F(1, 62) = 0.01$, $p = .908$, partial $\eta^2 = .000$, power = 0.05 and $F(1, 62) = 0.03$, $p = .867$, partial $\eta^2 = .000$, power = 0.05 respectively.

Due to the lack of statistically significant results among the knowledge test scores, an analysis was conducted on the knowledge test questions that assessed participants' ability to manually count sperm cells. Therefore, chi-square tests of independence were used to compare participants' absolute cell counts to expert counts to determine whether any observable patterns existed in mastery level based on treatment group for the immediate and delayed knowledge posttests. There are certain assumptions that must be met to confidently interpret the results of the chi-square test of independence (Gravetter & Wallnau, 2007).

The first condition is the requirement of two nominal variables. The first variable for this test was treatment group and the second variable was whether participants mastered the ability to manually count sperm cells. Correct answers were defined as within ±1 SD of the mean of expert counts. Mastery of the skill of performing manual sperm cell counts was determined by being judged correct on at least two out of three cell counts; whereas, participants who answered one or no count questions correctly were deemed to not have mastered the ability. Mastery was coded as 0 (not mastered) or 1 (mastered).

A second assumption for the chi-square analysis is that the observations are independent, meaning participants can only belong to one treatment group. Because participants were divided into four groups and only participated in one group, the assumption of independent observations was met.

The third assumption for the chi-square hypothesis test requires that the dependent variable be measured at a single time point. This study measured participants' ability to count sperm cells at three points in time (pre, immediate, and delayed); however, only the immediate posttest and delayed posttest results were examined and each were examined separately.

The last assumption for this analysis is that all cells should have expected frequencies greater than or equal to five. The immediate posttest showed that one cell (12.5%) had an expected count less than five and the delayed posttest showed that four cells (50%) had expected counts less than five; however, Glass and Hopkins (1984) indicated that the chi-square test works with an average expected frequency as low as two.

The results of the chi-square test of independence between treatment group and mastery level on the immediate posttest showed there was not a statistically significant association between treatment type and mastery level, $\chi^2(3) = 2.72$, $p = .437$. Similarly, there was not a statistically significant association between treatment type and mastery level for the delayed posttest, $\chi^2(3) = 2.20$, $p = .532$. The treatment conditions did not appear to have an effect on participants' accuracy for manually counting sperm cells. Table 25 shows the mastery results for all participants on the immediate and delayed knowledge posttests.

Table 25

Immediate and Delayed Posttest Mastery Levels by Treatment Group

	Immediate Posttest Mastery Level		Delayed Posttest Mastery Level	
Treatment Group	Mastered	Not Mastered	Mastered	Not Mastered
No scenario, elaborative (sE)	9 (69.23%)	4 (30.77%)	3 (23.08%)	10 (76.92%)
No scenario, intrinsic (sI)	9 (50.00%)	9 (50.00%)	3 (16.67%)	15 (83.33%)
Scenario, elaborative (SE)	9 (56.25%)	7 (43.75%)	2 (12.50%)	14 (87.50%)
Scenario, intrinsic (SI)	14 (73.68%)	5 (26.32%)	6 (31.58%)	13 (68.42%)
Total	41 (62.12%)	25 (37.88%)	14 (21.21%)	52 (78.79%)

Note. Chi-square analyses included all participants (N = 66).

It is worth noting that although a significant increase in knowledge scores from the immediate posttest to delayed posttest was observed, a decrease in mastery level from the immediate posttest to delayed posttest was also observed, as was to be expected. These results suggest that although knowledge appeared to improve over time, skill attainment was not mastered in a long-term sense. The largest decrease in mastery was observed for participants in the SE group where 69% of those participants met mastery on the immediate posttest, but only 23% demonstrated mastery on the delayed posttest, which was a decrease of 46%. It is also worth noting that the participants in the sE and SI groups demonstrated the highest mastery on the immediate posttest (69% and 74% respectively) and on the delayed posttest (23% and 32% respectively).

Research Question Two. Research Question Two examined whether the use of SBeL and feedback type have a significant effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's (2010) IMMS delivered as a pre, immediate post, and delayed post-measure. Research question two and the corresponding sub-parts as well as the null and alternate hypotheses are presented here.

2. Does the use of SBeL and feedback type have a significant effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's (2010) IMMS pretest, immediate posttest, and delayed posttest?

    a. Does the use of SBeL have a significant main effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's IMMS pretest, immediate posttest, and delayed posttest? The null hypothesis was that there would not be a difference in IMMS scores based on scenario level (H_0: $\mu_{\text{Scenario}} = \mu_{\text{No scenario}}$) and the alternate hypothesis was that a difference would exist between participants' IMMS scores based on scenario level (H_1: $\mu_{\text{Scenario}} \neq \mu_{\text{No scenario}}$).

    b. Does feedback type have a significant main effect over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on Keller's IMMS pretest, immediate posttest, and delayed posttest? The null hypothesis was that there would not be a difference in IMMS scores based on feedback type (H_0: $\mu_{\text{Elaborative}} = \mu_{\text{Intrinsic}}$) and the alternate hypothesis was that a difference would exist between participants' IMMS scores over time based on feedback type (H_1: $\mu_{\text{Elaborative}} \neq \mu_{\text{Intrinsic}}$).

    c. Is there a significant interaction effect of the combination of SBeL and feedback type over time on college students' motivation to manually count sperm cells as measured by the comparison of participant results on

> Keller's IMMS pretest, immediate posttest, and delayed posttest? The null hypothesis was that there would not be an interaction between scenario level and feedback type on the IMMS scores. The alternate hypothesis was that there would be an interaction effect between the two independent variables on IMMS scores.

For accurate inferences from this 2×2 repeated measures ANOVA, the conditions of a continuous dependent variable, one categorical within subjects factor with at least two levels, and at least one categorical between-subjects factor with two or more levels were met. Further, the dependent variable for Research Question Two was ratings on Keller's (2010) IMMS instrument, which met the assumption of a continuous dependent variable. The study also included the within-subjects factor of time with three levels (pre-, immediate post-, and delayed post-measures) and two categorical between-subjects factors (scenario level and feedback type) with two levels (present/not present and elaborative/intrinsic respectively) to meet these assumptions.

Two outliers were removed from the IMMS scores as identified by boxplots. No additional outliers were identified by boxplot after removing the two initial outliers. Shapiro-Wilk's test of normality indicated the scores were normally distributed ($p > .05$) and the assumptions of homogeneity of variances ($p > .05$) and covariances ($p > .05$) were met as indicated by Levene's and Box's M tests respectively. Mauchly's test of sphericity was violated, $\chi^2(2) = 8.71$, $p = .013$; therefore, the Greenhouse-Geisser estimates were used.

There was no statistically significant difference in IMMS scores over time based on scenario level, $F(1.76, 105.52) = 0.72$, $p = .474$, partial $\eta^2 = .012$. Similarly, there was no statistically significant difference in IMMS scores over time based on feedback type, $F(1.76, 105.52) = 0.42$, $p = .630$, partial $\eta^2 = .007$. Further, there was no statistically significant interaction between scenario level and feedback type over time on IMMS scores, $F(1.76, 105.52) = 0.18$, $p = .810$, partial $\eta^2 = .003$.

The main effect of time showed a statistically significant difference in IMMS scores, $F(1.76, 105.52) = 3.62$, $p = .036$, partial $\eta^2 = .057$. Pairwise comparisons without using a familywise correction showed a statistically significant decrease from the pre-IMMS ($M = 140.92$) to the immediate post-IMMS ($M = 134.17$, 95% CI, 0.98 to 12.51, $p = .023$); however, no statistically significant differences were observed between the immediate post-IMMS ($M = 134.17$) and the delayed post-IMMS ($M = 137.14$, 95% CI, -8.02 to 2.09, $p = .246$). Similarly, no significant differences were observed between the pre-IMMS ($M = 140.92$) and delayed post-IMMS ($M = 137.14$, 95% CI, -0.34 to 7.91, $p = .072$).

As with Research Question One, two 2×2 factorial ANOVAs were used to examine the differences between scenario levels and feedback types on the immediate and delayed IMMS posttests separately to check that power was not lost as part of the repeated measures ANOVA.

The IMMS pretest scores indicated there were no statistically significant differences of feedback type $F(1, 62) = .133$, $p = .716$, partial $\eta^2 = .002$, power = 0.06. The results also indicated there was no statistically significant effect for scenario level $F(1, 62) = .305$, $p = .583$, partial $\eta^2 = .005$, power = 0.09. In addition, there was not a statistically significant interaction between feedback type and scenario level on the pre-IMMS scores $F(1, 62) = .189$, $p = .666$, partial $\eta^2 = .003$, power = 0.07.

An examination of the boxplots for the immediate post-IMMS revealed one outlier for the SE group and one extreme outlier for the sE group; however, the outliers were not removed for analysis as removing outliers could bias the results (T. Peterson, personal communication, June 1, 2020). The data met the assumption for homogeneity of variances according to Levene's test for equality of variances ($p > .05$). The distributions of group scores met the assumption of normality according to Shapiro-Wilk's test of normality ($p > .05$) except for the sE group ($p = .039$); however, the ANOVA is fairly robust to violations of normality and violations of normality tend to lead to false positive results and no such positive results were observed. There was not a statistically significant interaction between feedback type and scenario level on the immediate post-IMMS, $F(1, 62) = 0.69$, $p = .408$, partial $\eta^2 = .011$, power = 0.13. The main effect of feedback type on the immediate post-IMMS was also not statistically significant, $F(1, 62) = 0.19$, $p = .664$, partial $\eta^2 = .003$, power = 0.07. Similarly, the main effect of scenario level on the immediate post-IMMS was not statistically significant, $F(1, 62) = 1.20$, $p = .279$, partial $\eta^2 = .019$, power = 0.20.

An examination of the boxplots for the delayed post-IMMS residuals revealed one outlier for the sI group that was not removed for analysis to avoid creating biased results. The Shapiro-Wilk's test indicated the residuals were normally distributed ($p > .05$) and Levene's test for equality of variances showed there was homogeneity of variances ($p > .05$). There was not a statistically significant interaction between feedback type and scenario level for the delayed post-IMMS, $F(1, 62) = 1.32$, $p = .256$, partial $\eta^2 = .021$, power = 0.22. Additionally, there were no statistically significant main effects for scenario level or feedback type, $F(1, 62) = 1.59$, $p = .212$, partial $\eta^2 = .025$, power = 0.25 and $F(1, 62) = 0.04$, $p = .835$, partial $\eta^2 = .001$, power = 0.06 respectively.

To determine whether there were any observable effects for the MLS or non-MLS participants, the data were divided into those subgroups and analyzed using 2×2 factorial ANOVAs. An examination of the boxplots for the MLS participants' immediate post-IMMS residuals revealed one extreme data point (more than three box-lengths away from the edge of the box) for the sE group that was not removed for analysis to avoid creating additional outliers or biased results. Subsequently, the residuals for the sE group were the only residuals that were not normally distributed as assessed by the Shapiro-Wilk's test of normality ($p < .05$). Levene's test for equality of variances showed there was homogeneity of variances ($p > .05$). There was not a statistically significant interaction between feedback type and scenario level for MLS participants' immediate post-IMMS ratings, $F(1, 41) = 0.00$, $p = .981$, partial $\eta^2 = .000$, power = 0.05. Additionally, there were no statistically significant main effects for scenario level or feedback type, $F(1, 41) = 0.80$, $p = .378$, partial $\eta^2 = .019$, power = 0.07 and $F(1, 41) = 0.01$, $p = .922$, partial $\eta^2 = .000$, power = 0.05 respectively.

An examination of the boxplots for the MLS participants' delayed post-IMMS residuals revealed no outliers. The assumption of normality was met as assessed by Shapiro-Wilk's test of normality ($p > .05$) and Levene's test for equality of variances showed there was homogeneity of variances ($p > .05$). There was not a statistically significant interaction between feedback type and scenario level for MLS participants' delayed post-IMMS ratings, $F(1, 41) = 0.59$, $p = .447$, partial $\eta^2 = .014$, power = 0.06. Additionally, there was not a statistically significant main effect for scenario level $F(1,$

41) = 0.74, p = .393, partial η^2 = .018, power = 0.07. Similarly, there was not a statistically significant main effect for feedback type $F(1, 41) = 0.70$, p = .408, partial η^2 = .017, power = 0.07.

An examination of the residual boxplots for the non-MLS participants' immediate post-IMMS residuals revealed one outlier for the sI group and one extreme outlier (more than three box-lengths away from the edge of the box) for the SE group. These data points were not removed for analysis to avoid creating additional outliers or biased results. The assumptions of normality and homogeneity of variances were met as indicated by the Shapiro-Wilk test of normality ($p > .05$) and Levene's test for equality of variances ($p > .05$) respectively. There was not a statistically significant interaction between feedback type and scenario level for non-MLS participants' immediate post-IMMS ratings, $F(1, 17) = 1.94$, p = .182, partial η^2 = .102, power = 0.10. Additionally, there were no statistically significant main effects for scenario level or feedback type for non-MLS participants' immediate post-IMMS ratings, $F(1, 17) = 0.45$, p = .512, partial η^2 = .026, power = 0.06 and $F(1, 17) = 0.29$, p = .595, partial η^2 = .017, power = 0.06 respectively.

One outlier was noted for the non-MLS participants in the SE group on the delayed post-IMMS as assessed by boxplots of the residuals, but was not removed to avoid creating additional outliers or biased results. The assumption of normality was met as assessed by Shapiro-Wilk's test of normality ($p > .05$) and Levene's test for equality of variances showed there was homogeneity of variances ($p > .05$). There was not a statistically significant interaction between feedback type and scenario level for non-MLS participants' delayed post-IMMS ratings, $F(1, 17) = 0.92$, p = .350, partial η^2 = .052, power = 0.07. Additionally, there was not a statistically significant main effect for scenario level $F(1, 17) = 0.59$, $p = .453$, partial $\eta^2 = .034$, power = 0.06 or for feedback type $F(1, 17) = 0.20$, $p = .662$, partial $\eta^2 = .012$, power = 0.05. The lack of statistically significant findings of scenario level and feedback type for the MLS and non-MLS participants indicates there were no treatment-based differences in motivation based on whether participants were enrolled in an MLS program.

Summary of the Results

The goal of this study was to examine the effects of SBeL and two feedback types on learner performance and motivation for manually counting sperm cells. The majority of participants in this study were females (64%) over 25 years old (61%) enrolled in an MLS program (60%). The bulk of the participants (93%) had previously taken an online course and had completed an online exam or assessment (95%). Additionally, 67% of the participants had previous experience with the LMS used for this study, but only 14% had received previous manual cell count instruction (as reported during the first week of each treatment iteration).

This research focused on two research questions. The first focused on the effects of scenario level and feedback type over time on learner performance as measured by a researcher-developed knowledge instrument. The second research question focused on the effects of scenario level and feedback type over time on participants' motivation as measured by Keller's (2010) IMMS.

The results of a 2×2 repeated measures ANOVA showed no statistically significant main effects of scenario level or feedback type over time as well as no statistically significant interaction between these variables on participants' knowledge as measured by a researcher developed knowledge instrument. Similarly, a second 2×2 repeated measures ANOVA showed no statistically significant main effects of scenario level or feedback type over time as well as no statistically significant interaction between these variables on participants' motivation or confidence as measured by Keller's (2010) IMMS. Therefore, there was not sufficient evidence to reject the null hypotheses for research question one or research question two.

A statistically significant main effect of time was observed for participants' scores between all three knowledge measurements, which could be interpreted as being a result of receiving the instructional treatment. There was also a statistically significant main effect of time for the IMMS ratings. Pairwise comparisons showed this difference to be a significant decrease from the pre-IMMS to the immediate post-IMMS, which could be interpreted as participants losing motivation after receiving the instructional treatments.

Chapter V

Discussion

The purpose of this study was to examine the efficacy of scenario-based e-learning (SBeL) and two types of feedback on learner performance and motivation. There were four treatment groups for this study and all instructional materials focused on a laboratory procedure for manually counting sperm cells. The intended population for this study was participants in Medical Laboratory Science (MLS) or Medical Laboratory Technologist (MLT) programs in the United States. The majority of participants (68%) were from a single MLS program at a medium-sized university in the Intermountain West while the bulk of the remaining participants (86%) were also from the same university. The researcher made three calls for participants; however, each call seemed to return diminishing numbers. Additionally, due to COVID-19, a voluntary study may not have been a priority for many participants. This chapter presents an interpretation of the results for each research question followed by recommendations for future research and future practice.

Interpretation of the Results

This study focused on two research questions. The interpretation of the results for each research question will be discussed below.

Research Question One. Research Question One asked if scenario level (present or not present) or feedback type (elaborative or intrinsic) had an effect on learning outcomes or if there was an interaction between these two variables. The results of the 2×2 repeated measures analysis of variance (ANOVA) showed no significant main effects over time for scenario level or feedback type. Additionally, no statistically significant interaction effect was observed between feedback type and scenario level over time. It is possible that the lack of statistical significance was due to the short treatment duration or small sample size. These findings suggest that participants' ability to manually count sperm cells was equivalent regardless of treatment condition.

Mean scores increased from the pretest to the immediate posttest for participants in all treatment groups, which was to be expected as a result of the instruction; however, mean scores also increased from the immediate posttest to the delayed posttest for participants in all groups, which was unexpected. These differences in mean knowledge scores were statistically significant. The observed increase from the pretest to the immediate posttest is attributed to having received the treatment instruction. Participants were not able to review the instructional interventions and were therefore unable to review the instruction prior to the delayed posttest. However, it is possible that the participants took notes and reviewed those notes as they completed the immediate posttest or the delayed posttest.

Another possible explanation for the increase between the pretest and immediate posttest is the additional manual white blood cell (WBC) and red blood cell (RBC) count instruction the MLS participants received in another course in the MLS program. Similarly, the increased scores between the immediate and delayed posttests may have been due to the lab activities some MLS participants may have completed prior to taking the delayed posttest. However, when the MLS and non-MLS participants' data were analyzed separately, the same pattern of increased mean scores from the immediate posttest to delayed posttest was observed for both MLS and non-MLS participants. The observed difference from immediate to delayed posttest for non-MLS participants was statistically significant (p = .013). It is highly unlikely the non-MLS participants received additional cell count instruction and they also were not able to review the instructional materials prior to taking the delayed posttest. It is possible non-MLS participants reviewed any notes they may have taken prior to completing the delayed posttest. Additionally, it is possible that participants completed the delayed posttest as an open-note test since the use of notes was not explicitly prohibited, nor was the use of notes monitored or prevented.

Although no significant differences were found between treatment conditions over time, patterns did seem to emerge. Figure 7 (repeated below) illustrates that the delayed posttest means improved more for participants who experienced no scenario with elaborative feedback (sE) compared to those who experienced no scenario with intrinsic feedback (sI).

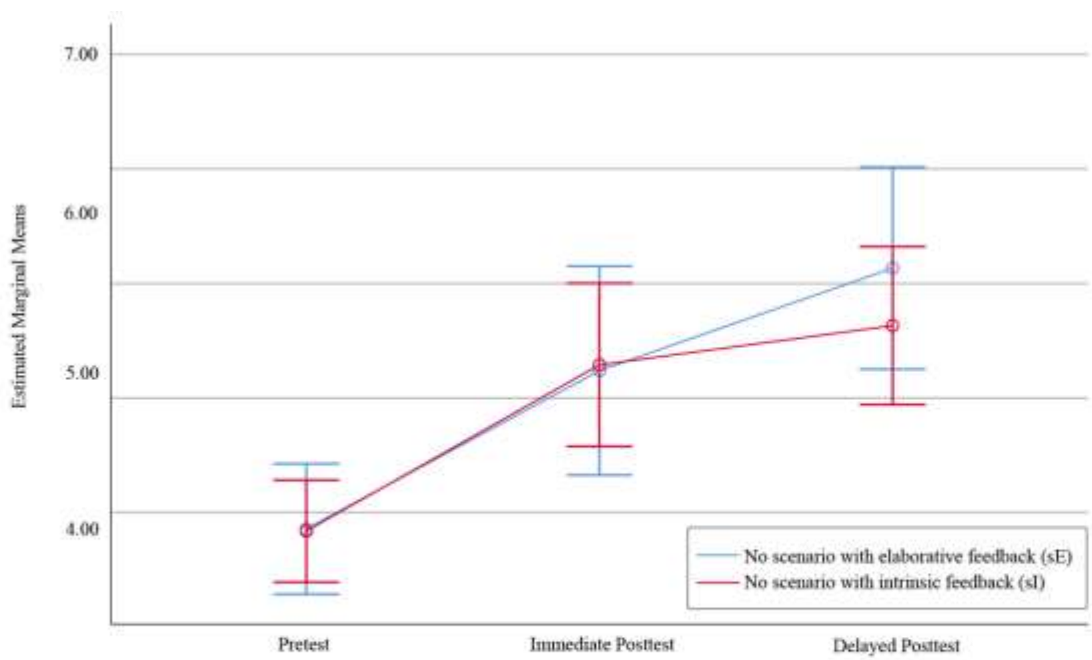


Figure 7. Comparison of knowledge scores by feedback type. Estimated marginal means for participants in the no scenario with elaborative feedback group (sE) and participants in the no scenario with intrinsic feedback group (sI). Error bars represent 95% Confidence Interval.

Similarly, participants in the scenario with intrinsic feedback (SI) group scored higher on the delayed posttest than those in the scenario with elaborative feedback (SE) group (see Figure 8 repeated below). These patterns suggest participants performed better on the delayed posttest when scenarios were paired with intrinsic feedback and when elaborative feedback was used without a scenario.

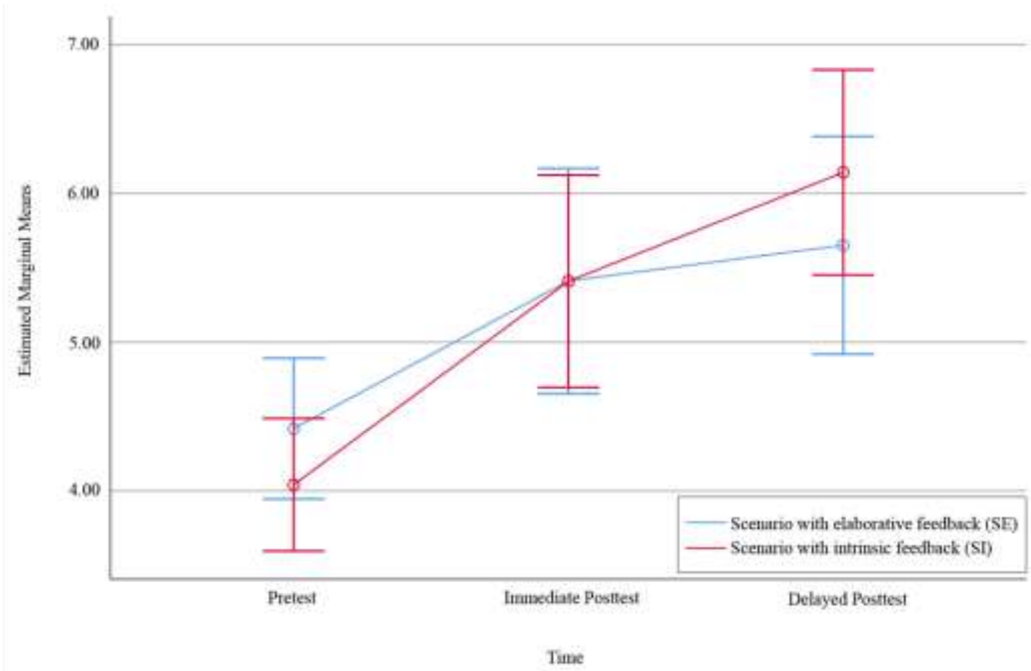


Figure 8. Comparison of knowledge scores by scenario level. Estimated marginal means for participants in the scenario with elaborative feedback group (SE) and participants in the scenario with intrinsic feedback group (SI). Error bars represent 95% Confidence Interval.

As was the case for all participants, the delayed posttest results also showed a higher mean score for MLS participants in the sE group than those in the SE group. Similarly, MLS participants in the SI group scored higher on the delayed posttest than MLS participants in the sI group. The non-MLS participants also exhibited this pattern on both the immediate and delayed posttests. These pairings might prove to be statistically significant over longer periods of time or with a larger sample size, which will be discussed later.

The chi-square test of independence for participants' ability to manually count sperm cells showed no significant differences between mastery levels based on treatment group for the immediate posttest or delayed posttest. These results indicate that all four treatments were equally effective for teaching participants how to manually count sperm cells. All treatments included one worked example and two practice activities. It is possible that this was enough instruction regardless of treatment condition for participants to be able to manually count sperm cells immediately after the instruction. Although no long-term mastery gains were observed for any of the treatment groups, the data did indicate that the highest mastery levels on the immediate posttest were among participants in the sE (69%) and SI (74%) groups. Similarly, 23% of participants in the sE group and 32% in the SI group achieved mastery on the delayed posttest. Again, this pattern seems to support the use of elaborative feedback with no scenario and intrinsic feedback with scenarios. Due to the small sample size, it was not possible to split this data into MLS and non-MLS participants to examine whether this pattern existed for those subgroups. Further, any observed pattern would need to be confirmed with a larger sample.

Although knowledge scores increased significantly between the immediate posttest and delayed posttest, mastery for manually counting sperm cells was not found to be significantly influenced by scenario level or feedback type. In fact, mastery decreased from the immediate posttest to the delayed posttest for all participants regardless of treatment group. The immediate and delayed knowledge instruments included 27 questions with the majority (89%) covering content rather than the ability to manually count sperm cells, which would support using online laboratory instruction for teaching conceptual content rather than practical skills. It is possible the knowledge scores increased because the majority of questions were content-based rather than skills-based; therefore, it is possible participants were able to demonstrate conceptual knowledge, but did not master the skill of manually counting sperm cells. Another possibility is that participants did not receive ample practice applying the practical skill. Both of these possibilities will be more fully discussed as part of the recommendations for future research below.

Research Question Two. Research Question Two asked if scenario level (present or not present) or feedback type (elaborative or intrinsic) had an effect on motivation or confidence or if there was an interaction between these two variables as measured by Keller's (2010) Instructional Materials Motivation Survey (IMMS). The results of the 2×2 repeated measures ANOVA showed no statistically significant main effects for scenario level or feedback type on motivation over time. Additionally, no statistically significant interaction effect between scenario level and feedback type over time on participants' motivation was observed. As with the knowledge instrument, the IMMS results showed a statistically significant main effect of time. Participants' IMMS ratings decreased significantly from the pre-IMMS to the immediate post-IMMS, which could have been due to participants being more interested in the study at the start, but having lost interest after receiving the instructional intervention. It is also possible that this decrease was due to the fact that participants felt they would not perform well on the immediate knowledge posttest.

A non-statistically significant pattern emerged that was consistent. Participants who received elaborative feedback rated their confidence higher from the pre-IMMS to the immediate post-IMMS, whereas participants who received intrinsic feedback rated their confidence lower from the pre-IMMS to the immediate post-IMMS. However, between the immediate post-IMMS and delayed post-IMMS measures, participants who received intrinsic feedback rated their confidence higher while those who received elaborative feedback rated their confidence lower. Due to the low power of this study, this pattern will be discussed later as a recommendation for future research.

Recommendations for Future Practice

It is possible the participants in this study did not fully attend to the feedback provided and as Hattie and Gan (2011) asserted, "Feedback not received is unlikely to have any effect on learning (p. 265); however, there has been compelling evidence in previous studies for the effect of feedback on learning outcomes (Johnson & Priest, 2014; Hattie & Gan, 2011; Shute, 2008) and Clark (2013) suggested that although feedback has the potential to improve learning, not all feedback is effective. The results of this study indicate that the elaborative and intrinsic feedback types may be freely interchanged without affecting participants' knowledge scores or motivation. This is in opposition to Hattie and Gan's (2011) assertion that, "There is a preponderance of evidence that feedback is a powerful influence in the development of learning outcomes (p. 249). However, as Hattie and Gan also pointed out, the substantial variance of effects indicates some types of feedback are more powerful than others. Clark and Mayer (2011) also suggested that including explanations with feedback capitalizes on the teaching moment of missed questions. Due to the low power of this study it may still be appropriate to recommend including elaborative or intrinsic feedback to provide guidance especially for novice learners.

Although, feedback has also been shown to influence motivation (Burkšaitienė, 2011), the results of this study seem to indicate that elaborative and intrinsic feedback did not impact motivation. Burkšaitienė (2011) suggested, that effective feedback can increase both student effort and motivation to foster learning; however, based on the results of this study, neither elaborative nor intrinsic feedback had an effect on participants' self-reported motivation. Therefore, using elaborative or intrinsic feedback in future practice does not appear to be effective or necessary for improving motivation.

Similarly, the use of scenarios did not statistically significantly impact participants' learning outcomes or self-reported motivation. The results of this study support Clark's (2016) assertion that, "SBeL is well suited for problem-solving tasks and for learners with some relevant background knowledge and skills" (p. 54). Based on the results of this study, it does not appear that designing scenario-based instruction is necessary for teaching college students to manually count sperm cells. Additionally, this study found no evidence to suggest that scenarios are necessary for increasing college students' motivation to manually count sperm cells.

All of the instructional interventions for this study incorporated worked examples, which Clark and Mayer (2011) suggested are beneficial for building procedural skills particularly during initial learning stages. Because all of the instructional interventions appeared to increase knowledge scores, future practice should incorporate worked examples as part of procedural instruction for novice learners.

Recommendations for Future Research

The purpose of this study was to examine the effects of scenarios and feedback type as well as any interaction between these variables on learning outcomes and motivation. Because the lack of statistically significant findings could have been due to the small sample size, future research using a larger sample is strongly recommended. This would require involvement of additional MLS and MLT programs, which would also help ensure the relevance of the real-world scenarios.

Descriptive statistics for Research Question One showed a potential for an interaction effect between scenario level and feedback type. For example, MLS (and the entire sample) participants in the SE group had the highest scores on the immediate posttest while the MLS (and entire sample) participants in the SI group had the highest mean score on the delayed knowledge posttest. However, non-MLS participants in the SI group had the highest mean score on the immediate knowledge posttest while non-MLS participants in the sE group had the highest mean score on the delayed knowledge posttest. Although these results were not statistically significant, this showed a pattern that the participants who were interested in the MLS career had more long-term gains when they received intrinsic feedback with scenarios and participants who were not interested in an MLS career had higher long-term gains when they received elaborative feedback without scenarios. Because this study's power to observe a statistically significant interaction effect was low (partial $\eta^2 = .010$, power = 0.13), future research is recommended to confirm any possible statistical significance of this pattern with a larger sample.

This study examined the efficacy of SBeL for teaching a specific type of manual cell counting. Future research should examine the efficacy of SBeL for teaching other types of laboratory procedures. Additionally, elaborative and intrinsic feedback types should be examined with other types of laboratory procedures. A qualitative element could also be added to future research to collect participants' reactions to the scenarios and feedback types.

It is also possible that the treatment was not an effective approach for teaching skill acquisition. The results of this study contradict Lim, Reiser, and Olina's (2009) findings that undergraduate pre-service teachers who were taught using a whole-task approach were better able to complete the whole task than those who were taught only partial tasks. Similarly, Lehmann, Bosse, Simon, Nikendei, and Huwendiek (2013) found that multimedia-enhanced virtual patients (VPs) were useful for providing preparatory activities for practical skills training. Clark (2016) suggested SBeL should be considered when "Building skills on the job would take too long or impose unacceptable risk" (p. 52); however, future research is needed to determine whether SBeL works better for teaching conceptual knowledge than for the skill of manually counting cells.

Additionally, because the content of the scenarios may not appeal to participants outside of MLS or MLT programs, future research limiting the participants to those enrolled in MLS or MLT programs is recommended. Further, the introduction of similar material in a different course could have influenced the results of this study; therefore, additional research is recommended with more control over content taught in other courses.

It is possible that the instructional intervention for this study did not include enough practice. Clark and Mayer (2011) argued that e-learning practice activities can be expensive to create and time consuming for learners. Therefore, they suggested these activities should, "strike a balance with practice assignments that require enough processing for learning but do not overload learners" (p. 258). Further, the authors suggested that evidence supports spacing practice over time. Future research is recommended to determine whether spaced repetition would be beneficial for teaching manual cell counts.

Similarly, it is possible the treatment duration for this study was too short to observe statistically significant effects. During week one of each study iteration, participants completed the knowledge pretest and pre-IMMS. Then during week four, participants completed the instructional intervention followed by the immediate posttest instruments. Participants spent an average of 90 minutes on the instruction followed by no further study activities until week seven. Future research could examine whether a longer treatment duration has an effect on learning outcomes or motivation. Additionally, future research could examine whether time spent on task has an effect on learning outcomes or motivation.

Although there is extensive feedback research, the results have been varied and inconsistent (Shute, 2008). As Shute suggested, "Care should be taken to know which interventions increase performance and under which conditions" (p. 170). Therefore, future research could examine whether elaborative or intrinsic feedback types are more effective for certain learners or for certain learning outcomes. Further, if the patterns observed in this study persist with larger samples, future research could examine whether elaborative feedback works best without scenarios and whether intrinsic feedback works better with scenarios for both conceptual knowledge and skills acquisition.

Lehmann et al. (2013) examined the use of case studies with multimedia-enhanced VPs to prepare medical students prior to skills labs in an effort to have more time during labs to spend on practice and feedback rather than instruction. Based on their results, the authors concluded that multimedia-enhanced VPs were useful preparatory activities for practical skills training. Future research could examine SBeL's effectiveness for preparing students prior to classroom laboratory sessions.

Clark (2013) suggested, "We need more evidence regarding the optimal modes, that is, text, video, animation, or stills to represent scenario-based e-learning cases" (p. 151). Based on the results of this study, future research could examine whether making the scenarios more realistic using audio, video, or other multimedia might have more of an effect on learning outcomes or motivation. For example, future research could deliver the scenarios using patient videos similar to Lehman et al.'s (2013) VPs to determine whether scenarios that are more realistic have an effect on learning outcomes or motivation.

A study with more power might find statistical significance for the consistent pattern that emerged for participants' confidence ratings on the IMMS measure. The non-statistically significant drop of participants' confidence ratings from the pre-IMMS to the immediate post-IMMS suggests that participants' who received intrinsic feedback had decreased confidence in their ability to learn after receiving the instruction, but then gained some confidence back for the delayed post-IMMS. On the contrary, participants who received elaborative feedback had a boost in confidence from the pre-IMMS to the immediate post-IMMS, but then dropped that confidence for the delayed post-IMMS. Therefore, future research is recommended to examine the potential for elaborative or intrinsic feedback to have an effect on learners' confidence.

Conclusions

The results of this study add to the body of knowledge on SBeL and feedback as well as motivational design for learning. Clark (2016) noted that SBeL is not necessarily appropriate for all learners or all learning outcomes; however, the results of this study showed that SBeL did not appear to impede learning, confidence, or motivation. Based on the results of this study, both intrinsic and elaborative feedback appear to be comparable for teaching college students how to manually count sperm cells as well as motivating them and building their confidence to do so.

Participants in all treatment groups appeared to increase their knowledge for manually counting sperm cells regardless of the scenario level or feedback type. Further, the treatment conditions did not significantly impact motivation or confidence. Therefore, the results of this study support Clark's (2013) assertion that SBeL could be considered for instances when in-person training is rare, dangerous, or impractical; however, more research is needed to determine if the results of this study hold true with larger sample sizes or for different learners and learning outcomes.

References

Allen, I. E., & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States*. Sloan Consortium. PO Box 1238, Newburyport, MA 01950.

Allen, I. E., & Seaman, J. (2016). *Online report card: Tracking online education in the United States*. Babson Survey Research Group. Babson Park, MA.

Allen, M. W. (2016). *Michael Allen's guide to e-learning: Building interactive, fun, and effective learning programs for any company*. Hoboken, NJ: John Wiley & Sons.

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, *13*(2), 111-127.

Bozorgmanesh, M., Sadighi, M., & Nazarpour, M. (2011). Online education: Definitions and advantages in adult education. *Researcher*, *3*(4), 74-78.

Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. *American Educational Research Journal*, *5*(4), 437-474.

Breakey, K. M., Levin, D., Miller, I., & Hentges, K. E. (2008). The use of scenario-based-learning interactive software to create custom virtual laboratory scenarios for teaching genetics. *Genetics*, *179*(3), 1151-1155.

Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *92*(1), 81-90.

Brinson, J. R. (2015). Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research. *Computers & Education*, *87*, 218-237.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32-42.

Bruning, R. H., Schraw, G. J., & Ronning, R. R. (1999). *Cognitive psychology and instruction*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Burkšaitienė, N. (2011). Promoting student learning through feedback in higher education. *Societal Studies*, *4*(1), 33-46.

Campbell, D. T., & Stanley, J.C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.

Carnevale, D. (2003). The virtual lab experiment. *Chronicle of Higher Education*, *49*(21), A30.

Carpenter, K. M., Watson, J. M., Raffety, B., & Chabal, C. (2003). Teaching brief interventions for smoking cessation via an interactive computer-based tutorial. *Journal of Health Psychology*, *8*(1), 149-160.

Carroll, J. M. (1999, January). Five reasons for scenario-based design. In *Systems Sciences, 1999. HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on* (pp. 1-11). Institute of Electrical and Electronics Engineers.

Casey, D. M. (2008). The historical development of distance education through technology. *TechTrends*, *52*(2), 45.

Centers for Medicare & Medicaid Services. (2018). Retrieved from https://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/Program_Descriptions_Projects.html

Choi, J. I., & Hannafin, M. (1995). Situated cognition and learning environments: Roles, structures, and implications for design. *Educational Technology Research and Development*, *43*(2), 53-69.

Choules, A. P. (2007). The use of elearning in medical education: A review of the current situation. *Postgraduate Medical Journal*, *83*(978), 212-216.

Clariana, R. B., Wagner, D., & Murphy, L. C. R. (2000). Applying a connectionist description of feedback timing. *Educational Technology Research and Development*, *48*(3), 5-22.

Clark, R. C. (2013). *Scenario-based e-learning: Evidence-based guidelines for online workforce learning*. San Francisco, CA: John Wiley & Sons.

Clark, R. C. (2014). Multimedia learning in e-courses. In Mayer, R. (Ed.), *The Cambridge handbook of multimedia learning* (pp. 842-881). New York, NY: Cambridge University Press.

Clark, R. C. (2016). Accelerate expertise with scenario-based e-learning. *TD: Talent Development*, *70*(8), 50. Retrieved from https://www.td.org/magazines/accelerate-expertise-with-scenario-based-e-learning

Clark, R. C., & Mayer, R. E. (2011). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. San Francisco, CA: John Wiley & Sons.

Collins, A. (1991). Cognitive apprenticeship and instructional technology. *Educational values and cognitive instruction: Implications for reform*, 121-138.

Cook, D. A., Levinson, A. J., Garside, S., Dupras, D. M., Erwin, P. J., & Montori, V. M. (2008). Internet-based learning in the health professions: a meta-analysis. *Journal of the American Medical Association*, *300*(10), 1181-1196.

Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago, IL: Rand McNally.

David, L. (2007, January 17). Situated learning theory (Lave). Retrieved from https://www.learning-theories.com/situated-learning-theory-lave.html

Errington, E. P. (2010). Preparing graduates for the professions: achieving employability through the exploration of near-world scenarios. *International Journal of Interdisciplinary Social Sciences*, *5*, 1-10.

Errington, E. P. (2011). Mission possible: Using near-world scenarios to prepare graduates for the professions. *International Journal of Teaching and Learning in Higher Education*, *23*, 84-91.

Ertmer, P. A., Richardson, J. C., Belland, B., Camin, D., Connolly, P., Coulthard, G. . . . & Mong, C. (2007). Using peer feedback to enhance the quality of student online postings: An exploratory study. *Journal of Computer-Mediated Communication*, *12*(2), 412-433.

Galbraith, K. J. (2011). *Effects of student-sequenced learning objects and student-selected multimedia formats on motivation and achievement in a family finances learning unit* (Doctoral dissertation). Idaho State University, Pocatello. Retrieved from https://www.learntechlib.org/p/120416/

Gibson, J. W. (2008). A comparison of student outcomes and student satisfaction in three MBA human resource management classes based on traditional vs. online learning. *Journal of College Teaching and Learning*, *5*(8), 1-9.

Glass, G. V., & Hopkins, K. D. (1984). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Gravetter, F. J., & Wallnau, L. B. (2007). *Statistics for the behavioral sciences.* Wadsworth, Cengage Learning.

Hart, C. (2012). Factors associated with student persistence in an online program of study: A review of the literature. *Journal of Interactive Online Learning*, *11*(1), 19-42.

Hattie, J., & Gan, M. (2011). Instruction based on feedback. In R. Mayer, P. Alexander (Eds.), *Handbook of research on learning and instruction*, (pp. 249-271). New York, NY: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112.

Herrington, J., & Oliver, R. (1997). Multimedia, magic and the way students respond to a situated learning environment. *Australian Journal of Educational Technology*, *13*(2), 127-143.

Herrington, J., & Oliver, R. (1999). Using situated learning and multimedia to investigate higher-order thinking. *Journal of Interactive Learning Research*, *10*(1), 3-24.

Herrington, J., Oliver, R., & Reeves, T. C. (2003). Patterns of engagement in authentic online learning environments. *Australian Journal of Educational Technology*, *19*(1), 59-71.

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*(1), 28-54.

Hursen, C., & Fasli, F. G. (2017). Investigating the Efficiency of Scenario Based Learning and Reflective Learning Approaches in Teacher Education. *European Journal of Contemporary Education*, *6*(2), 264-279.

Jaehnig, W., & Miller, M. L. (2007). Feedback types in programmed instruction: A systematic review. *The Psychological Record*, *57*(2), 219-232.

Johnson, C. I., & Priest, H. A. (2014). The feedback principle in multimedia learning. In Mayer, R. (Ed.), *The Cambridge handbook of multimedia learning* (pp. 449-463). New York, NY: Cambridge University Press.

Jones, R. (2018). The need for balanced feedback. Retrieved from https://www.facultyfocus.com/articles/online-education/the-need-for-balanced-feedback/

Kee, N., Matthews, S., & Perumalla, C. (2009). Lessons learned from an online human physiology course. *Proceedings of the International Conference on e-Learning*, 247-254.

Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. New York, NY: Springer Science & Business Media.

Kentnor, H. E. (2015). Distance education and the evolution of online learning in the United States. *Curriculum and Teaching Dialogue*, *17*(1), 21-34.

Kiechle, F. L. (2017). Q&A. *CAP TODAY*, *31*(2), 64.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, *41*(2), 75-86.

Landrigan, B. (2010). Using online scenario-based learning modules to assist NESB students studying Corporations Law. In *ASCILITE-Australian Society for Computers in Learning in Tertiary Education Annual Conference* (Vol. 2010, No. 1, pp. 523-527).

Lechermeier, J., & Fassnacht, M. (2018). How do performance feedback characteristics influence recipients' reactions? A state-of-the-art review on feedback source, timing, and valence effects. *Management Review Quarterly*, *68*(2), 145-193.

Lee, H. S., & Butler, N. (2003). Making authentic science accessible to students. *International Journal of Science Education*, *25*(8), 923-948.

Lehmann, R., Bosse, H. M., Simon, A., Nikendei, C., & Huwendiek, S. (2013). An innovative blended learning approach using virtual patients as preparation for skills laboratory training: perceptions of students and tutors. *BMC Medical Education*, *13*(1), 1.

Lim, J., Reiser, R. A., & Olina, Z. (2009). The effects of part-task and whole-task instructional approaches on acquisition and transfer of a complex cognitive skill. *Educational Technology Research and Development*, *57*(1), 61-77.

Loyens, S. M., Jones, S. H., Mikkers, J., & van Gog, T. (2015). Problem-based learning as a facilitator of conceptual change. *Learning and Instruction*, *38*, 34-42.

Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In Mayer, R. (Ed.), *The Cambridge handbook of multimedia learning* (pp. 43-71). New York, NY: Cambridge University Press.

Mayer, R. E. (2014b). Principles based on social cues. In Mayer, R. (Ed.), *The Cambridge handbook of multimedia learning* (pp. 345-368). New York, NY: Cambridge University Press.

Mayer, R. E. (2014c). Research-based principles for designing multimedia instruction. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* (pp. 59-70). Retrieved from the Society for the Teaching of Psychology web site http://teachpsych.org/ebooks/asle2014/index.php

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of educational psychology*, *95*(4), 806-812. doi: 10.1037/0022-0663.95.4.806

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In Mayer, R. (Ed.), *The Cambridge handbook of multimedia learning* (pp. 279-315). New York, NY: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*(1), 107-119.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist, 38*(1), 43-52.

Medical Technology Schools website. (n.d.). Retrieved from https://www.medicaltechnologyschools.com/medical-lab-scientist/mlt-vs-mls

Meisner, G., Hoffman, H., & Turner, M. (2008). Learning physics in a virtual environment: Is there any? *Latin-American Journal of Physics Education*, *2*(2), 2.

Meldrum, K. (2011). Preparing pre-service physical education teachers for uncertain future(s): A scenario-based learning case study from Australia. *Physical Education and Sport Pedagogy*, *16*(2), 133-144.

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same? *The Internet and Higher Education*, *14*(2), 129-135.

Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional science*, *32*(1-2), 99-113.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*(3), 309-326.

Morrison, G. R., Ross, S. M., Kemp, J. E., & Kalman, H. (2011). *Designing effective instruction*. Hoboken, NJ: John Wiley & Sons.

Myers, J. L., Well, A. D., Lorch Jr., R. F. (2010). *Research design and statistical analysis (3rd Ed.)*. New York, NY: Routledge.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199-218.

Oliver, R., & Herrington, J. (2000). Using situated learning as a design strategy for Web-based learning. In *Instructional and cognitive impacts of web-based education* (pp. 178-191). IGI Global.

Park, J. H., & Choi, H. J. (2009). Factors influencing adult learners' decision to drop out or persist in online learning. *Journal of Educational Technology & Society*, *12*(4), 207-217.

Renkl, A. (2014). The worked examples principle in multimedia learning. In Mayer, R. (Ed.), *The Cambridge handbook of multimedia learning* (pp. 391-412). New York, NY: Cambridge University Press.

Rothmann, S. A., & Reese, A. A. (2007). Semen analysis: The test techs love to hate. *Medical Laboratory Observer*, *39*(4), 18.

Rumelhart, D., & Ortony, A. (1977). The representation of knowledge in memory. In Anderson, R., Spiro, R., & Montague, W. (Eds.), *Schooling and the acquisition of knowledge* (pp. 99–135). Hillsdale, NJ: Lawrence Erlbaum.

Sener, J. (2005). Escaping the comparison trap: Evaluating online learning on its own terms. *Innovate: Journal of Online Education*, *1*(2), 5.

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

Sorin, R. (2013). Scenario-based learning: transforming tertiary teaching and learning. In *Proceedings of the 8th QS-APPLE Conference, Bali* (pp. 71-81). James Cook University.

Stonebraker, I., Robertshaw, M. B., & Moss, J. D. (2016). Student see versus student do: A comparative study of two online tutorials. *TechTrends*, *60*(2), 176-182.

Stufflebeam, D. L., Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass.

Sung, H. Y., Hwang, G. J., & Yen, Y. F. (2015). Development of a contextual decision-making game for improving students' learning performance in a health education course. *Computers & Education*, *82*, 179-190.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295-312.

Sweller, J. (1999). *Instructional design in technical areas. Australian Education Review, No. 43*. Camberwell, Victoria: The Australian Council for Educational Research Ltd.

Sweller, J., Van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251-296.

The World Bank. (2002). *Constructing knowledge societies: New challenges for tertiary education. A World Bank Report.* Retrieved from http://siteresources.worldbank.org/TERTIARYEDUCATION/Resources/Documents/Constructing-Knowledge-Societies/ConstructingKnowledgeSocieties.pdf

Tichavsky, L. P., Hunt, A. N., Driscoll, A., & Jicha, K. (2015). "It's just nice having a real teacher": Student perceptions of online versus face-to-face instruction. *International Journal for the Scholarship of Teaching and Learning*, *9*(2), 2.

Toth, E.E., Morrow, B., & Ludvico, L. (2009). Designing blended inquiry learning in a laboratory context: A study of incorporating hands-on and virtual laboratories. *Innovative Higher Education*, *33*(5), 333-344. doi:10.1007/s10755-008-9087-7

Tzu-Chien, L. (2005). Web-based cognitive apprenticeship model for improving pre-service teachers' performances and attitudes towards instructional planning: Design and field experiment. *Educational Technology & Society, 8*(2), 136-149.

Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research*, *85*(4), 475-511.

Van der Kleij, F. M., Timmers, C. F., & Eggen, T. J. (2011). The effectiveness of methods for providing written feedback through a computer-based assessment for learning: A systematic review. *Cadmo*. 21-38. doi:10.3280/CAD2011-001004

Van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In Mayer, R. (Ed.), *The Cambridge handbook of multimedia learning* (pp. 263-278). New York, NY: Cambridge University Press.

Van Merriënboer, J., & Kester, L. (2014). The four-component instructional design model: Multimedia principles in environments for complex learning. In Mayer, R. (Ed.), *The Cambridge handbook of multimedia learning* (pp. 104-148). New York, NY: Cambridge University Press.

Waldrop, M. M. (2013). The virtual lab. *Nature, 499*, 268-270.

Xu, X., Kauer, S., & Tupy, S. (2016). Multiple-choice questions: Tips for optimizing assessment in-seat and online. *Scholarship of Teaching and Learning in Psychology*, *2*(2), 147.

Yarnall, L., Toyama, Y., Gong, B., Ayers, C., & Ostrander, J. (2007). Adapting scenario-based curriculum materials to community college technical courses. *Community College Journal of Research and Practice, 31*(7), 583-601.

Yetik, S. S., Akyuz, H. I., & Keser, H. (2012). Preservice teachers' perceptions about their problem solving skills in the scenario based blended learning environment. *Turkish Online Journal of Distance Education, 13*(2), 158-168.

Appendix A

Permissions to Reprint/Adapt Figures

PARTIES:

1. **Cambridge University Press** [CompanyNumber] (Licensor); and
2. **Sacha Johnson** (Licensee).

Thank you for your recent permission request. Some permission requests for use of material published by the Licensor, such as this one, are now being facilitated by PLSclear.

Set out in this licence cover sheet (the **Licence Cover Sheet**) are the principal terms under which Licensor has agreed to license certain Licensed Material (as defined below) to Licensee. The terms in this Licence Cover Sheet are subject to the attached General Terms and Conditions, which together with this Licence Cover Sheet constitute the licence agreement (the **Licence**) between Licensor and Licensee as regards the Licensed Material. The terms set out in this Licence Cover Sheet take precedence over any conflicting provision in the General Terms and Conditions.

Free Of Charge Licence Terms

Licence Date:	25/03/2019
PLSclear Ref No:	11595

The Licensor

Company name:	Cambridge University Press
Address:	University Printing House Shaftesbury Road Cambridge CB2 8BS GB

The Licensee

Licensee Contact Name:	Sacha Johnson
Licensee Address:	605 Wayne Ave. Pocatello 83201 United States

Licensed Material

title:	Multimedia Learning
ISBN:	9780521787499
publisher:	Cambridge University Press

figure number & title / caption	Figure 3.3 (A) The auditory/verbal channel (Top Frame) and (B) visual/pictorial channel (Bottom Frame) in a cognitive theory of multimedia learning.
Are you requesting permission to reuse your own work?	No. I am NOT the author
page number	47
position on page	entire page
Are you using the content as a prop?	content will NOT be used as a prop
additional information	I would like to use an adaptation of this figure in my doctoral dissertation, which will be published.
reproduction colour	Black and White
will it be cropped	Yes
full details of how it will be altered	I will create my own figure using the upper portion of the image ("Top Frame"), but I will not highlight the auditory/verbal channel. I'll also use Times New Roman font as that is the font required for my dissertation. The figure will be reproduced i

For Use In Licensee's Publication(s)

usage type	Book, Journal, Magazine or Academic Paper...-Thesis
distribution	online PDF, bound book at library and department
estimated publication date	May, 2020
language	English
number of pages	102
other relevant Information	I'm currently writing my research proposal so the page numbers will change once my research has finished. The title is yet to be determined as well; however, my research centers on scenario-based e-learning and is informed by the cognitive theory of
publication title	TBD
type of document	Doctoral dissertation

Rights Granted

Exclusivity:	Non-Exclusive
Format:	Thesis
Language:	English
Territory:	
Duration:	Lifetime of Licensee's Edition
Maximum Circulation:	0

Idaho State UNIVERSITY

Sacha Johnson <sevesach@isu.edu>

Permission to use figure

Richard Mayer <mayer@ucsb.edu> Thu, Feb 21, 2019 at 10:40 PM
To: Sacha Johnson <sevesach@isu.edu>

You have me permission to use the figure in your dissertation. Best wishes, Rich Mayer

> On Feb 21, 2019, at 5:05 PM, Sacha Johnson <sevesach@isu.edu> wrote:
>
> Good afternoon Drs. Moreno and Mayer,
>
> I am writing to request your permission to reprint Figure 1 from your 2007 Interactive Multimodal Learning Environments article for my doctoral dissertation, which will be published. Please let me know if this is acceptable. Alternatively, I'm willing to build my own image, but would still need your permission to recreate the figure.
>
> Thank you for your time and consideration,
>
> Sacha Johnson
> Instructional Design doctoral candidate, ABD
> 208.282.3954
> Senior Instructional Technologist
> Schedule an appointment
> Instructional Technology Resource Center
> Idaho State UNIVERSITY
>
> *Notice: This e-mail, including attachments, is covered by the Electronic Communications Privacy Act, 18 U.S.C. §§ 2510-2521, is confidential and may be legally privileged. If you are not the intended recipient, you are hereby notified that any retention, dissemination, distribution, or copying of this communication is strictly prohibited. Please reply to the sender (only) that you have received the message in error, then delete it. Unauthorized interception of the e-mail is a violation of federal criminal law. Thank you!*

Richard E. Mayer
Distinguished Professor
Department of Psychological and Brain Sciences
University of California
Santa Barbara, CA 93106

Confirmation Number: 11803360
Order Date: 03/30/2019

Customer Information

Customer: Sacha Johnson
Account Number: 3001306136
Organization: Sacha Johnson
Email: sevesach@isu.edu
Phone: +1 (208) 380-1644
Payment Method: Invoice

This is not an invoice

Order Details

Motivational Design for Learning and Performance

Billing Status: **N/A**

Order detail ID:	71865121
ISBN:	978-1-4419-1250-3
Publication Type:	Book
Publisher:	Springer
Author/Editor:	Keller, John M.

Permission Status: **Granted**

Permission type: Republish or display content
Type of use: Thesis/Dissertation
Order License Id: 4558981107173

Requestor type	Academic institution
Format	Print, Electronic
Portion	chart/graph/table/figure
Number of charts/graphs/tables /figures	1
The requesting person/organization	Sacha Johnson / Idaho State University
Title or numeric reference of the portion(s)	(adaptation of) Chapter 1, Figure 1.2
Title of the article or chapter the portion is from	The Study of Motivation
Editor of portion(s)	N/A
Author of portion(s)	John M. Keller
Volume of serial or monograph	N/A
Page range of portion	6
Publication date of portion	2010
Rights for	Main Product, any product related to main product, and other compilations/derivative products
Duration of use	Life of current and all future editions
Creation of copies for the disabled	no
With minor editing privileges	yes
For distribution to	Worldwide
In the following language(s)	Original language of publication
With incidental promotional use	yes
Lifetime unit quantity of new product	Up to 499
Title	(TBD) examining the effect of scenario-based e-learning and two types of feedback on learning outcomes and motivation
Institution name	Idaho State University
Expected presentation date	May 2020

Note: This item was invoiced separately through our **RightsLink service.** More info **$ 0.00**

Total order items: 1 **Order Total: $0.00**

Appendix B

Permission to use Instructional Materials Motivation Survey

Idaho State UNIVERSITY

Sacha Johnson <sevesach@isu.edu>

Permission to use IMMS

John Keller <jkellersan@gmail.com> Tue, Apr 16, 2019 at 2:11 PM
To: Sacha Johnson <sevesach@isu.edu>

Dear Sacha,

Thank you for your polite request! You certainly are welcome to use the IMMS in your research. Your projected research sounds interesting and I wish you success in your results!

In case you don't have it, I am attaching a section of my book which contains the instrument and information about it.

Best wishes and feel free to write again if you have any questions.
Sincerely,
John

On Tue, Apr 16, 2019 at 1:24 PM Sacha Johnson <sevesach@isu.edu> wrote:

Good morning Dr. Keller,

I am an (ABD) instructional design doctoral candidate at Idaho State University and I am writing to request your permission to use your Instructional Materials Motivation Survey instrument for my dissertation study. I plan to research the effects of scenario-based e-learning and two types of feedback on learning outcomes and motivation. I will be proposing my research on May 2nd and will conduct my study this coming fall. Please let me know at your earliest convenience if I may use your IMMS instrument for my study. You're welcome to call or email me if you have any questions.

Thank you for your time and consideration,

Sacha Johnson
208.380.1644 (cell)
208.282.3954 (office)
Senior Instructional Technologist
Schedule an appointment
Instructional Technology Resource Center
Idaho State UNIVERSITY

Notice: This e-mail, including attachments, is covered by the Electronic Communications Privacy Act, 18 U.S.C. §§ 2510-2521, is confidential and may be legally privileged. If you are not the intended recipient, you are hereby notified that any retention, dissemination, distribution, or copying of this communication is strictly prohibited. Please reply to the sender (only) that you have received the message in error, then delete it. Unauthorized interception of the e-mail is a violation of federal criminal law. Thank you!

Keller 2010 ARCS Measurement Surveys.pdf
4219K

Appendix C

Scenario and No-scenario Lesson Comparison

Scenario-based Treatment Lesson Example Pages

Export as PDF

Manually Counting Sperm Cells

First day on the job

Congratulations! You've just landed your first dream job in a medical laboratory. One of the first tests that have come in is a semen analysis. Jill and Jack Parkes have been trying over three years to conceive and have decided to pursue fertility testing as a result of their unsuccessful efforts. Because of its relative simplicity, semen examination is often requested as part of an infertility investigation before the more complicated and expensive examination of the female so that's where their fertility doctor has decided to start. Male factors contribute to a significant percentage of infertility problems, estimated to be greater than 50 percent by some investigators. An abbreviated test, known as a post-vasectomy semen analysis, is also conducted after a vasectomy to confirm the absence of any motile sperm cells.

Your supervisor, Dave, noticed three patients scheduled for semen analysis today. He's decided to bring you in on this lab test as part of your training and competency to perform semen analysis in this lab. He shows you the first test requisition.

MLS Medical Laboratory
Idaho State University
Pocatello, Idaho 83209
Phone (208) 282-5765
Dr. QC M.D.
Laboratory Director

Name: Parkes, Jack
DOB: 10-6-84
Sex: Male
Diagnosis: N46.9 Male infertility
Requesting Physician: Stanley
Test Requested: Semen analysis
Specimen Source: Semen
Date and Time of Specimen Collection: 2-3-20 at 0800 by patient
Date and Time Specimen Received: 2-3-20 at 0810 by djr

Although Dave is going to perform the other portions of the test, he wants you to learn to perform the cell count. To start, he's instructed you to read the lab's standard operating procedure (SOP). He showed you the shelf with all of the binders of SOPs so you could find the Semen Analysis SOP.

After you've read the SOP click on the Start button below.

Start

Lesson menu

- First day on the job
- The Hemocytometer
- The Hemocytometer Under the Microscope
- Charging the Hemocytometer
- Counting Sperm Cells
- Training
- Show what you know

Export as PDF

Manually Counting Sperm Cells

Charging the Hemocytometer

Now that you've read the SOP and Dave has explained the dimensions of the hemocytometer, Dave walks you through the process of charging the hemocytometer, which means adding the sample to the chambers. He helped you make a dilution of Jack's sample where you added 19 parts diluent to 1 part semen, which is a 1:20 dilution. But during this process, he cautioned you that sometimes a different dilution (such as 1:5) is needed.

Dave explained that you'll need to charge the disposable hemocytometer by adding 10μL of the diluted specimen to each chamber and has set the pipette to 10μL, but explained that not all hemocytometers are designed to be charged with exactly 10μL and must be filled via capillary action until full. After adding the sample to the hemocytometer, Dave said you need to let the hemocytometer stand for 5-10 minutes to allow the cells to settle before examining them under the microscope.

Dave demonstrated the process of charging the hemocytometer for you as shown in the video below. Complete any interactions that appear during the video and submit your answers at the end of the video. Use the full screen button to enlarge the video and the Escape button to close the full screen. After Dave demonstrates this for you, click the Next button under the video to continue.

Previous Next

Lesson menu

Export as PDF

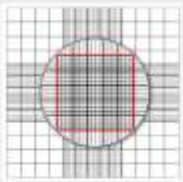
Manually Counting Sperm Cells

Training

As part of your training, Dave is going to demonstrate performing a cell count using Jack's sample. You can use the full screen button below to view the demonstration in a larger window and then use the Escape button to minimize the window again. After completing the demonstration, click on the Next button at the bottom of the page to see what Dave will have you do next.

Dave's diagram (below) shows the nine large squares of one counting chamber and the circled area shows the large center square, which he said is the area of interest for counting sperm cells. The image (at right) shows the hemocytometer with Jack's semen specimen in the large center square using the 10x objective.

Dave is going to walk you through the sperm cell count for Jack's semen analysis. Click anywhere on the 10x objective view to get started.



Previous Next

Lesson menu

- First day on the job
- The Hemocytometer
- The Hemocytometer Under the Microscope
- Charging the Hemocytometer
- Counting Sperm Cells
- Training
- Show what you know

Export as PDF

Manually Counting Sperm Cells

Now that Dave has demonstrated cell counting and you have the steps down, in order to be deemed trained and competent and independently report patient results, he wants to work another sample with you. It just so happens that Christine and Tim Matthews have been instructed to have a semen analysis conducted and Tim's sample came in this afternoon.

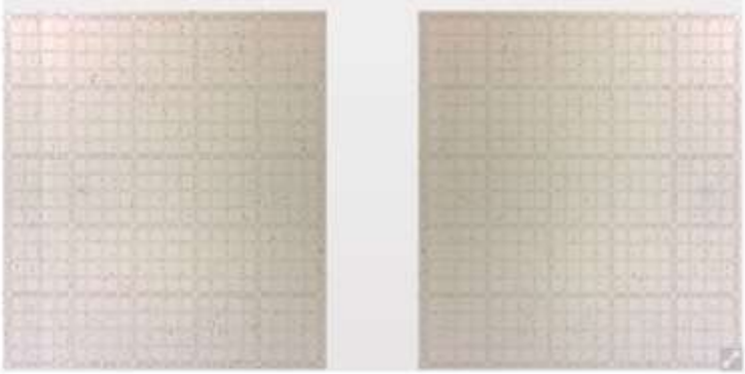
MLS Medical Laboratory
Idaho State University
Pocatello, Idaho 83209
Phone (208) 282-5765
Dr. QC M.D.
Laboratory Director

Name: Matthews, Tim
DOB: 2-4-80
Sex: Male
Diagnosis: N46.9 Male Infertility
Requesting Physician: Stanley
Test Requested: Semen analysis
Specimen Source: Semen
Date and Time of Specimen Collection: 2-3-20 at 1300 by patient
Date and Time Specimen Received: 2-3-20 at 1310 by dr

Since you're still fairly new at this, Dave wants you to perform the count portion of Tim's semen analysis while he completes the other portions. He's asked you to count the four small squares on each side as opposed to the five he had you count earlier and he'll provide feedback after you report your count to him.

The image below shows the large center squares on both sides of a standard Neubauer hemocytometer charged with Tim's diluted specimen and viewed using the 10x objective. Click on the four small squares on each side to count the cells using the 40x objective. Write down the number of cells in each small square because you'll use them to calculate the absolute count, which you'll report here. Considering that Dave measured the total volume to be 3.5 mL and you made a 1:20 dilution of this sample, what is the absolute total number of sperm cells in the specimen? *Enter only the whole numerical value without comma(s) or decimal points and then click the Submit button to show your results to Dave.*

Side A   Side B

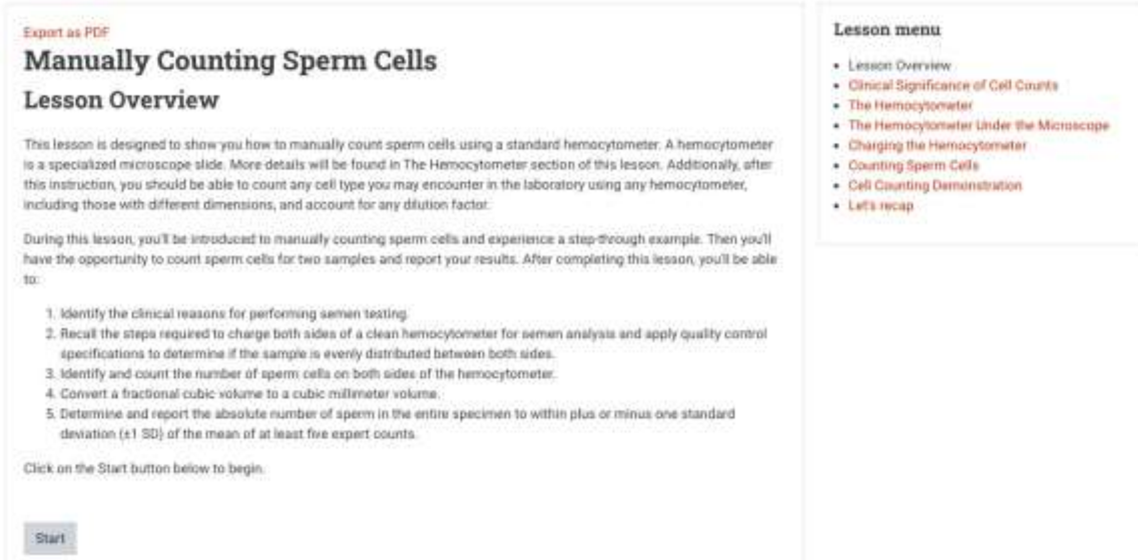


Your answer

Submit

Lesson menu

No Scenario-based Treatment Example Lesson Pages



Export as PDF

Manually Counting Sperm Cells

Lesson Overview

This lesson is designed to show you how to manually count sperm cells using a standard hemocytometer. A hemocytometer is a specialized microscope slide. More details will be found in The Hemocytometer section of this lesson. Additionally, after this instruction, you should be able to count any cell type you may encounter in the laboratory using any hemocytometer, including those with different dimensions, and account for any dilution factor.

During this lesson, you'll be introduced to manually counting sperm cells and experience a step-through example. Then you'll have the opportunity to count sperm cells for two samples and report your results. After completing this lesson, you'll be able to:

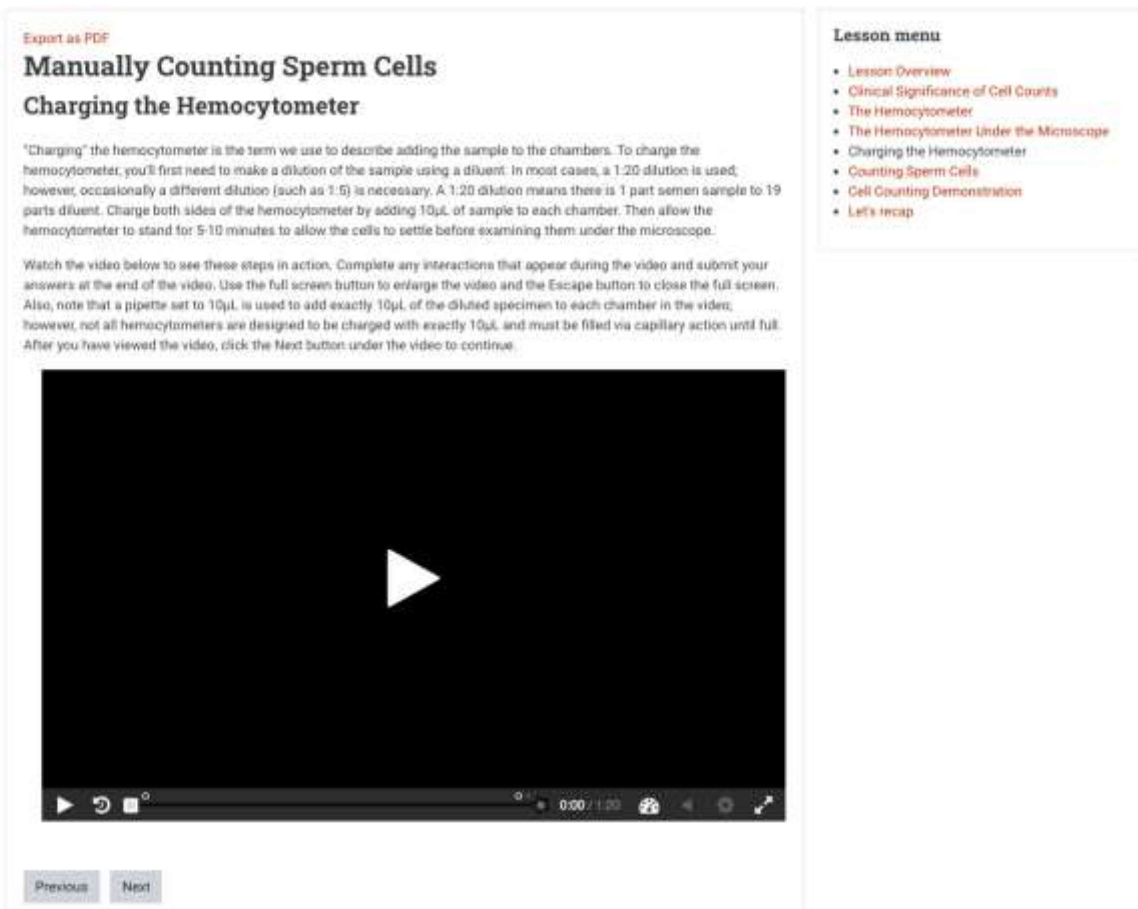
1. Identify the clinical reasons for performing semen testing.
2. Recall the steps required to charge both sides of a clean hemocytometer for semen analysis and apply quality control specifications to determine if the sample is evenly distributed between both sides.
3. Identify and count the number of sperm cells on both sides of the hemocytometer.
4. Convert a fractional cubic volume to a cubic millimeter volume.
5. Determine and report the absolute number of sperm in the entire specimen to within plus or minus one standard deviation (±1 SD) of the mean of at least five expert counts.

Click on the Start button below to begin.

Start

Lesson menu

- Lesson Overview
- Clinical Significance of Cell Counts
- The Hemocytometer
- The Hemocytometer Under the Microscope
- Charging the Hemocytometer
- Counting Sperm Cells
- Cell Counting Demonstration
- Let's recap



Export as PDF

Manually Counting Sperm Cells

Charging the Hemocytometer

"Charging" the hemocytometer is the term we use to describe adding the sample to the chambers. To charge the hemocytometer, you'll first need to make a dilution of the sample using a diluent. In most cases, a 1:20 dilution is used; however, occasionally a different dilution (such as 1:5) is necessary. A 1:20 dilution means there is 1 part semen sample to 19 parts diluent. Charge both sides of the hemocytometer by adding 10μL of sample to each chamber. Then allow the hemocytometer to stand for 5-10 minutes to allow the cells to settle before examining them under the microscope.

Watch the video below to see these steps in action. Complete any interactions that appear during the video and submit your answers at the end of the video. Use the full screen button to enlarge the video and the Escape button to close the full screen. Also, note that a pipette set to 10μL is used to add exactly 10μL of the diluted specimen to each chamber in the video; however, not all hemocytometers are designed to be charged with exactly 10μL and must be filled via capillary action until full. After you have viewed the video, click the Next button under the video to continue.

Previous Next

Lesson menu

- Lesson Overview
- Clinical Significance of Cell Counts
- The Hemocytometer
- The Hemocytometer Under the Microscope
- Charging the Hemocytometer
- Counting Sperm Cells
- Cell Counting Demonstration
- Let's recap

Export as PDF

Manually Counting Sperm Cells

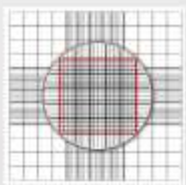
Cell Counting Demonstration

The following presentation provides a demonstration of the steps for calculating the absolute sperm cell count. You can use the full screen button to view the demonstration in a larger window and then use the Escape button to minimize the demonstration again. After you complete the demonstration, click on the Next button at the bottom of this page to continue with this lesson.

The diagram (below) shows the nine large squares of one counting chamber and the circled area shows the large center square, which is our area of interest when counting sperm cells.

The image (at right) shows an actual semen specimen in the large center square as viewed using the 10x objective.

Click anywhere on the 10x objective view of the semen sample to start this cell counting demonstration.



Previous Next

Lesson menu

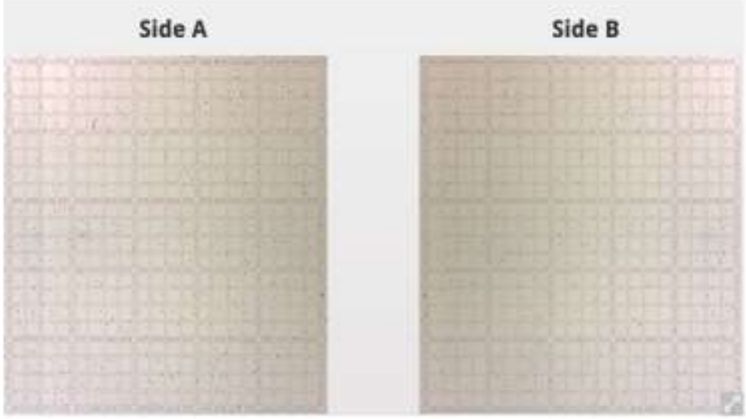
- Lesson Overview
- Clinical Significance of Cell Counts
- The Hemocytometer
- The Hemocytometer Under the Microscope
- Charging the Hemocytometer
- Counting Sperm Cells
- Cell Counting Demonstration
- Let's recap

Export as PDF

Manually Counting Sperm Cells

Now that you've gone through the cell counting demo and have the steps down, here's the first of two practice activities. In this practice, you'll count the four small squares on each side as opposed to the five we counted in the demonstration.

The image below shows the large center squares on both sides of a standard Neubauer hemocytometer using the 10x objective. Click on the four small squares on each side of the small center squares to count the cells using the 40x objective. Write down the number of cells in each small square because you'll use them to calculate the absolute count, which you'll report here. Considering that this is a 1:20 dilution of a 3.5 mL specimen, what is the absolute total number of sperm cells in the specimen? *Enter only the whole numerical value without comma(s) or decimal points and then click the Submit button.*



Your answer

Submit

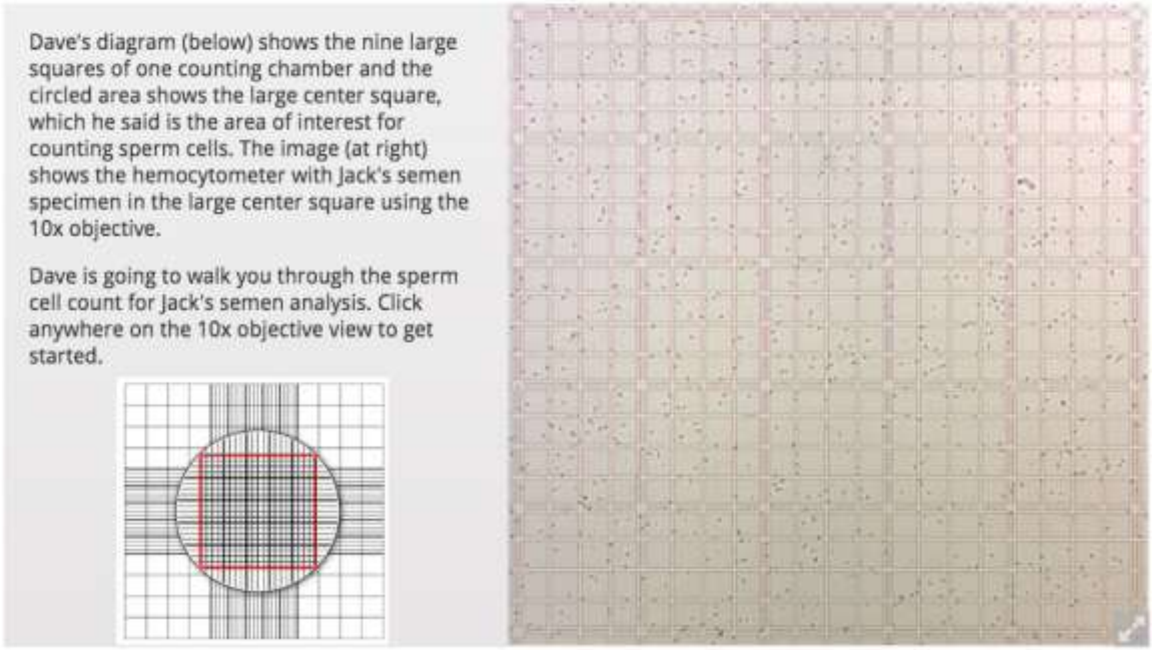
Lesson menu

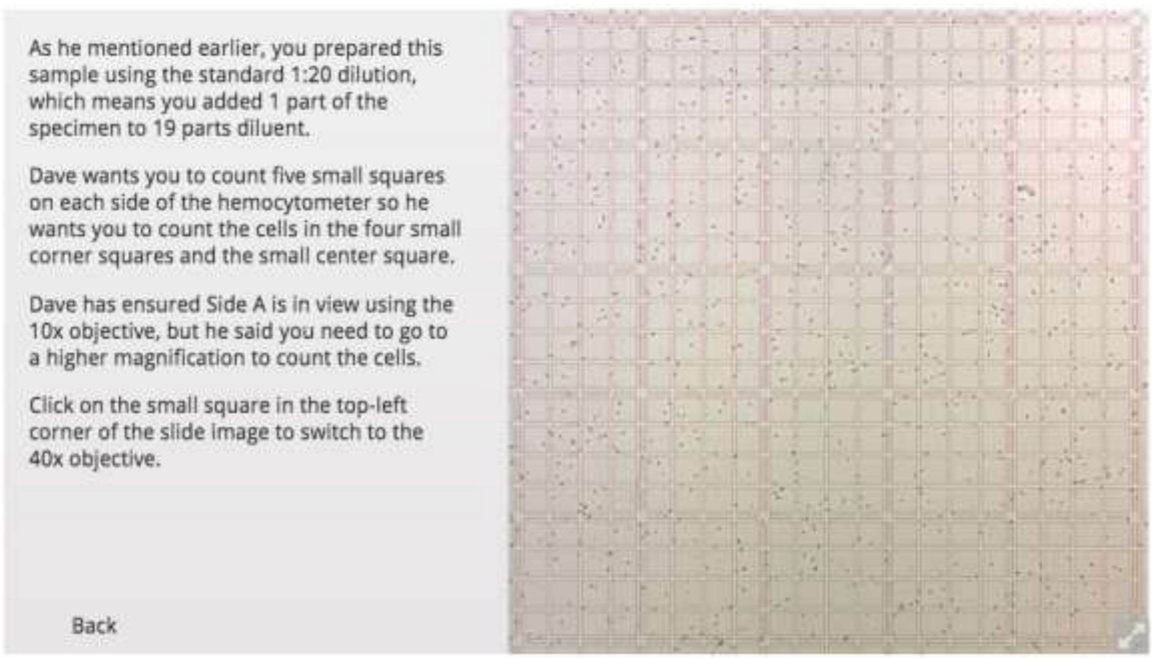
- Lesson Overview
- Clinical Significance of Cell Counts
- The Hemocytometer
- The Hemocytometer Under the Microscope
- Charging the Hemocytometer
- Counting Sperm Cells
- Cell Counting Demonstration
- Let's recap

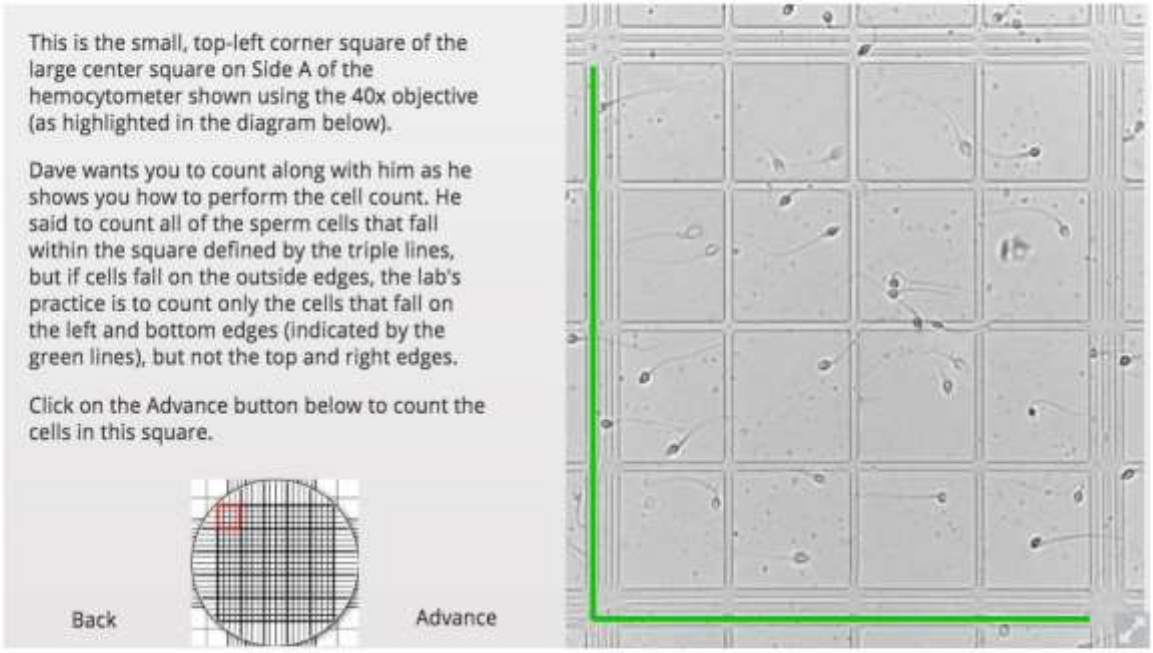
Appendix D

Scenario and No-scenario Worked Example Comparison

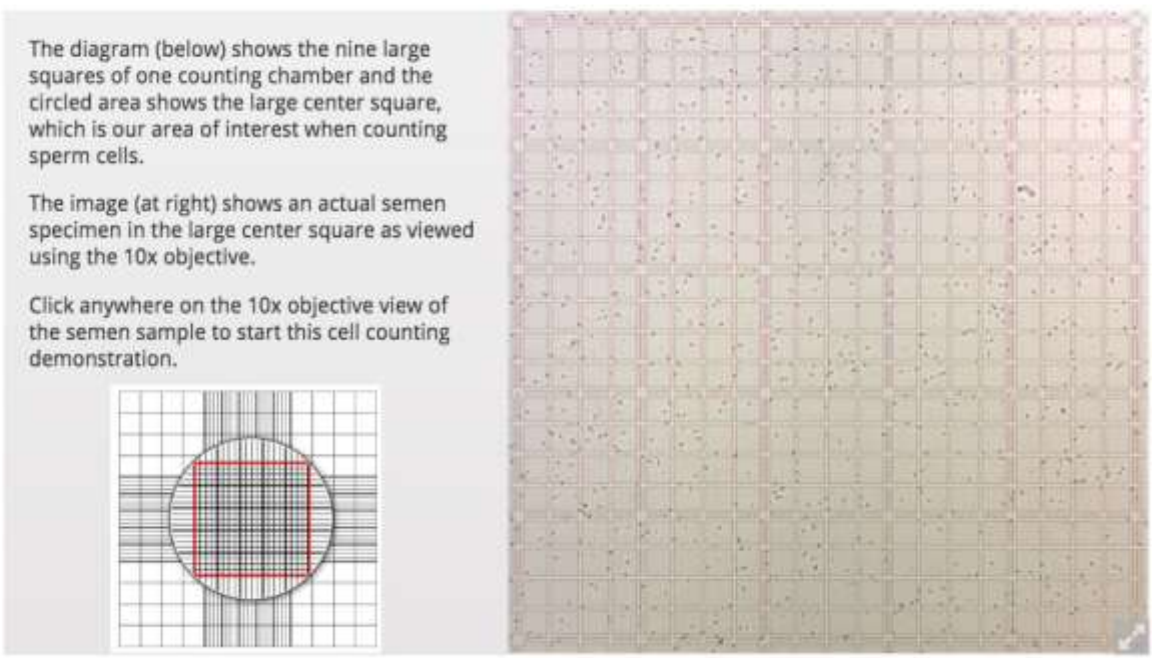
Scenario-based Worked Example







No Scenario Worked Example



In this demonstration, we're using the standard dilution of 1:20, which means 1 part specimen added to 19 parts diluent.

We're also counting five small squares on each side of the hemocytometer so we'll examine the four small corner squares and small center square.

You can see the large center square of Side A using the 10x objective; however, in order to count the cells, you need to go to a higher magnification.

Let's start our count by clicking on the small square in the top-left corner of the slide image to switch to the 40x objective.

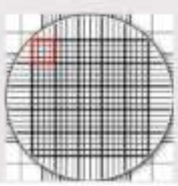
Back

This is the small, top-left corner square of the large center square on Side A of the hemocytometer shown using the 40x objective (as highlighted in the diagram below).

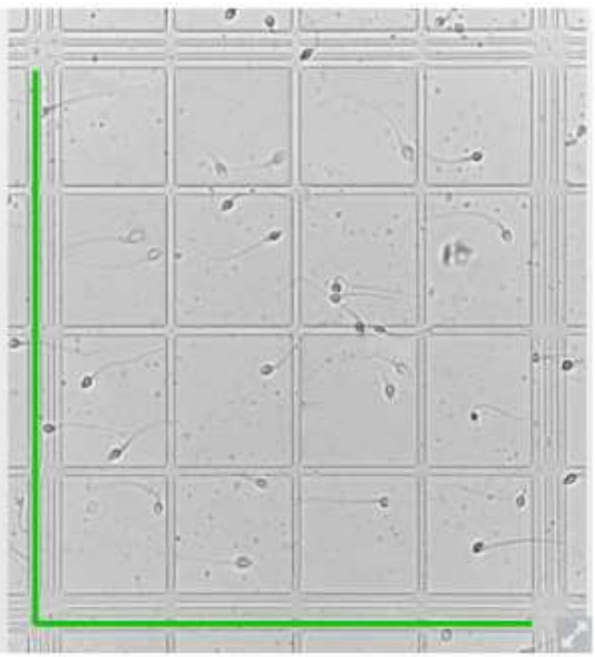
When counting cells in the hemocytometer, count all of the sperm in the square as defined by the triple lines. Count only the sperm cells that fall on two out of the four edges. In this module, we'll count any cells that fall on the left and bottom edges (indicated by the green lines), but not the top and right edges.

Click on the Advance button below to count the cells in this square.

Back

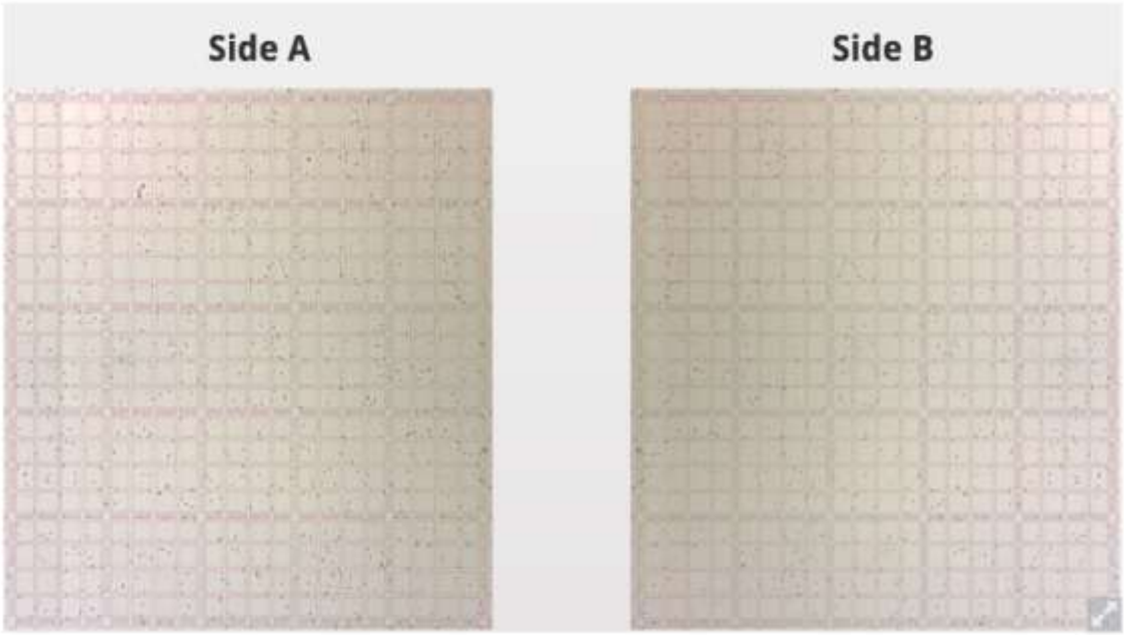


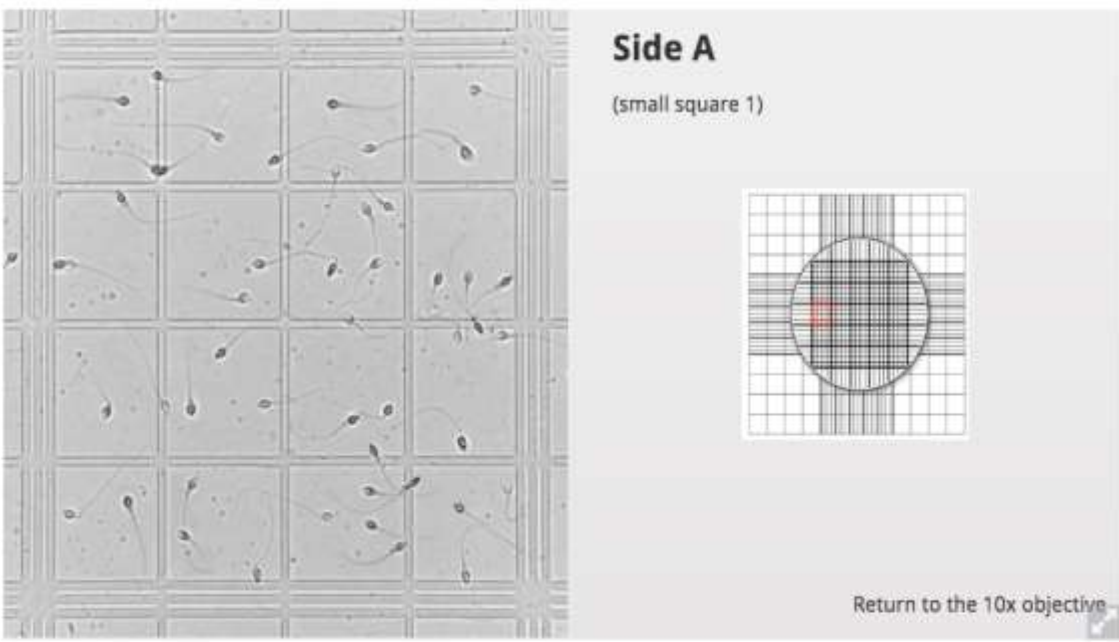
Advance

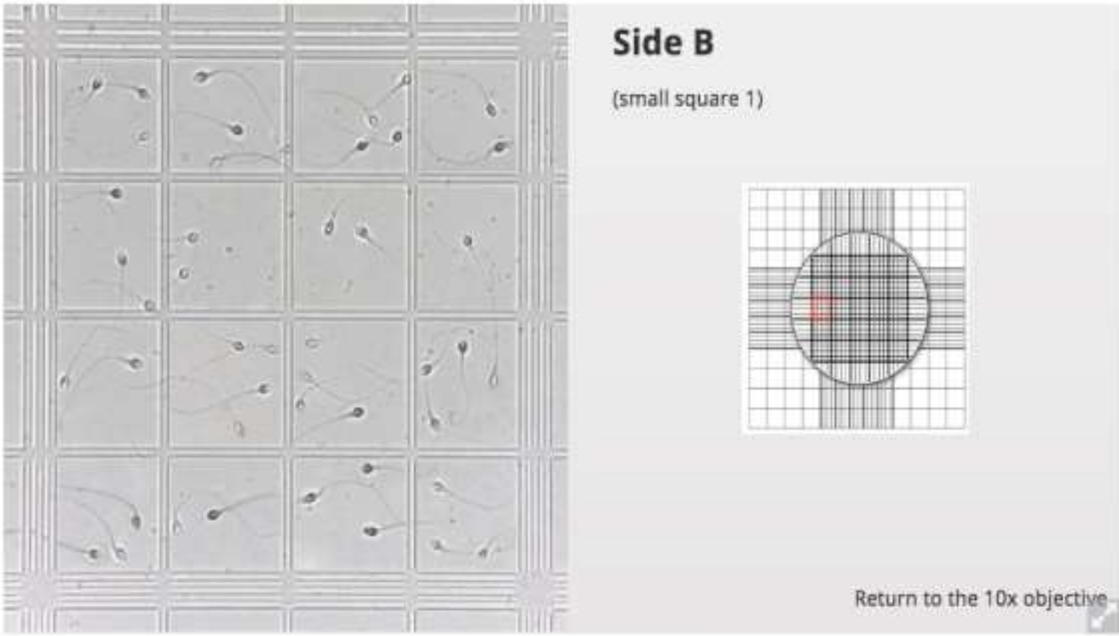


Appendix E

Practice Activity Screenshots


Side A
Side B


Side A
(small square 1)
Return to the 10x objective


Side B
(small square 1)
Return to the 10x objective

Appendix F

Intrinsic and Elaborative Feedback Comparison

Intrinsic Feedback Example

Response:
Although it was within the reference range, your count was significantly different from Dave's calculation of 857,500,000. Such an erroneous calculation could have negative clinical consequences. For example, depending on the other semen analysis parameters, mistakenly under-calculating the count parameter could lead the clinician to inform a couple that they're unlikely to conceive naturally based on erroneous information, which may cause them undue stress. Additionally, based on erroneous information, which may cause them undue stress. Additionally, the clinician may also suggest more costly and extensive fertility testing and procedures. Furthermore, believing a diagnosis of infertility could affect a and extensive fertility testing and procedures. Furthermore, believing a diagnosis of infertility could affect a patient's self-patient's self-esteem and lead to feelings of disappointment and inadequacy.

Dave's morphology and motility results were borderline, so it's especially important in this case that your count is accurate because of these clinical consequences. A discrepancy such as this typically indicates that a calculation error was made; however, Dave says clerical errors can happen so even if you calculated correctly it's possible to enter the value incorrectly into the computer and that it's important to consider whether the number you're reporting makes sense by referring to the normal reference range. He also said to make sure you double-check your answers before reporting them and then shows you what he calculated.

Although it was within the reference range, your count was significantly different from Dave's calculation of 857,500,000. Such an erroneous calculation could have negative clinical consequences. For example, depending on the other semen analysis parameters, mistakenly under-calculating the count parameter could lead the clinician to inform a couple that they're unlikely to conceive naturally based on erroneous information, which may cause them undue stress. Additionally, the clinician may also suggest more costly and extensive fertility testing and procedures. Furthermore, believing a diagnosis of infertility could affect a patient's self-esteem and lead to feelings of disappointment and inadequacy.

Dave's morphology and motility results were borderline, so it's especially important in this case that your count is accurate because of these clinical consequences. A discrepancy such as this typically indicates that a calculation error was made; however, Dave says clerical errors can happen so even if you calculated correctly it's possible to enter the value incorrectly into the computer and that it's important to consider whether the number you're reporting makes sense by referring to the normal reference range. He also said to make sure you double-check your answers before reporting them and then shows you what he calculated.

Elaborative Feedback Example

Response:
Incorrect, the correct answer is 857,500,000. It looks like you may have calculated incorrectly. We only counted 8 small squares in this practice so rather than dividing by 10 small squares, you would divide by 8. Additionally, the dilution factor was 4 and the total ejaculate volume was 2.0 mL. Something else to consider when reporting any lab results is whether the number you're reporting makes sense. In the case of sperm cell counts, be sure to refer to the normal reference range, which is greater than 60,000,000 cells.

Make sure you're using the correct values and following the proper steps as shown below to see where you may have miscalculated.

Incorrect, the correct answer is 857,500,000. It looks like you may have calculated incorrectly. We only counted 8 small squares in this practice so rather than dividing by 10 small squares, you would divide by 8. Additionally, the dilution factor was 4 and the total ejaculate volume was 2.0 mL. Something else to consider when reporting any lab results is whether the number you're reporting makes sense. In the case of sperm cell counts, be sure to refer to the normal reference range, which is greater than 60,000,000 cells.

Make sure you're using the correct values and following the proper steps as shown below to see where you may have miscalculated.