Effects of Interactive Online Instruction in Science Writing on the Development of Critical Analysis Skills in College Science Students

by

Cynthia Chesworth Adams

A Dissertation

presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

In

Learning Sciences and Technology

Lehigh University

April, 2015

CERTIFICATE OF APPROVAL

PROPOSED CANDIDATE: Cynthia Chesworth Adams, M.S.

PROGRAM CONCENTRATION: Learning Sciences & Technology

DATE APPROVED: ______________________

PROPOSAL COMMITTEE:

____

H. Lynn Columba, Ph.D., Chair
Associate Professor, Instructional
Technology
Lehigh University

____

Thomas C. Hammond, Ph.D.
Assistant Professor, Instructional
Technology
Lehigh University

____

Keith J. Schray, Ph.D.
Professor of Chemistry
Lehigh University

____

Warren R. Heydenberk, Ph.D.
Professor Emeritus, College of Education
Lehigh University

*Signatures are on file at Lehigh University in the College of Education.

Dedication

This thesis is dedicated to

My father,

E. Thomas Chesworth, Ph.D.,

Who proved by his example that

Graduate school is possible at almost any age.

and

To my children

Caitlin Hope and Erin Elizabeth

Love you always

Acknowledgements

My journey at Lehigh University has been the epitome of mastery learning, Benjamin Bloom's contention that anyone can learn if given enough time. In the almost nine years of my doctoral studies, I have learned the language of education and the transformative potential of technology. For that I must thank all of my professors, but especially Dr. Thomas Hammond, Dr. Scott Garrigan and Dr. M.J. Bishop who each patiently nurtured my meager technology skills through at least two courses. My original goal of becoming a better science teacher was accomplished thanks to what I learned from Dr. Lynn Columba, Dr. Ward Cates and Dr. Alec Bodzin. And I would be remiss if I did not thank Dr. Jill Sperandio, Dr. Gary Lutz and Dr. Grace Caskie for helping to hone my abilities to conduct and understand the results of educational research.

I would also like to acknowledge the patient support I have received throughout my years at Lehigh from the following technology team members Denise Campion, Bobby Siegfried, Dr. Judd Hark and Ilena Key.

I would also like to acknowledge and thank several people who played key roles in my dissertation project, especially Dr. Krystle McLaughlin of the Lehigh University Biological Sciences Department. Thank you for allowing your students to serve as subjects in my study, especially during your first semester of teaching at Lehigh University. Thank you also to Dr. Robert Skibbens, Dr. Lynn Cassimersis and Dr. Neal Simon of the same department who shared their expertise and obvious enthusiasm for science by "starring" in the video. I must also thank my fellow graduate student, Daniel Pennebacker and his high school class of computer programming students for creating the "Scientific Proof" game; you did a great job. Thanks also to graduate student Nasreen Haddush for Captivate advice that helped me create the tutorial.

I also owe a debt of gratitude to Dr. Precie Schroyer and Dr. Gail Mrowinski for the countless hours they spent refining the CA rubric and rating student papers: you are amazing. Thanks also to Dr. Pat Waller for rating more papers at the eleventh hour and to Dr. Qioung Fu for her patient assistance in helping me complete final statistical analyses.

I would also like to acknowledge the encouragement of the members of my doctoral committee, Dr. Thomas Hammond, Dr. Warren Heydenberk and Dr. Keith Schray and thank you all for your patience and flexibility throughout my dissertation study. Last but not least, I would like to thank my advisor, Dr. Lynn Columba-Piervallo; I could not have made it through this program without you.

On a personal note, I would also like to acknowledge Dr. Maria Kitchens-Kintz for her unfailing friendship and sage advice, as well as Todd Brunstetter for his persistent support and patient encouragement during the last few years of my Lehigh journey.

Table of Contents

Copyright ……………………………………………………………….ii

Certificate of Approval ……………………………………………….iii

Dedication ……………………………………………………………….iv

Acknowledgements ……………..…………………………………….v

Table of Contents ……………………………………………………..vi

List of Tables …………………………………………………………vii

List of Figures ………………………………………………………viii

Abstract ……………………………………………………….……..… 1

Chapter 1: Introduction…………………………………………………2

Definition of Terms……………………………………………………17

Chapter 2: Review of the Literature ……………………………….. 20

Chapter 3: Methodology……………………………………………….51

Chapter 4: Results……………………………………………………...74

Chapter 5: Discussion…………………………………………………104

References……………………………………………………………..115

Appendices……………………………………………………………142

List of Tables

Table 1 Study Treatment Groups ……..…………………………………...… 83

Table 2 Average Overall CA …………………………………………..……. 89

Table 3 IRR of key CA elements using Cronbach's Alpha ……...………. 90

Table 4 ANOVA results for Group Effects ……………………………….92

Table 5 ANOVA results for Paper Order Effects……………………….....93

Table 6 ANOVA results for Gender Effects ……………………………....93

Table 7 Intervention Log Entries…………………………………………..94

Table 8 Comparison of Pre- and Post-Test CA Ranking ………………….98

Table 9 Mean Student Ranking of CA Study Characteristics ……………..99

List of Figures

Figure 1 Study timeline……………………………………………………...…51

Figure 2 Sample scene from video……………………………………….….58

Figure 3 Sample scene from game…...…………………………….……….....60

Figure 4 Sample introductory page for video…………………………………..61

Figure 5 Pre-test: Word cloud – Where data found in articles.....……..………..…75

Figure 6 Pre-test: Compilation of answers about references …………………......75

Figure 7 Pre-test: Student ranking of own CA skills……...………………………76

Figure 8 Post-Test: Word cloud - Factors of Credible Scientific Literature ……...96

Figure 9 Pre/Post-Test: Comparison of Student CA Self-Ranking………………..97

Figure 10 Post-Test: Student Ranking of CA Study Characteristics ……………....99

Abstract

Everyone needs better critical analysis (CA) skills to evaluate and process the overflow of information available today. These skills are only going to be more crucial for the next generation given the logarithmic growth of factual information projected for the future (Çavdar & Doe, 2012; Webber, Boon & Johnston, 2005).

In this study, early-year college science analyzed two scientific papers in a counter-balance study. One analysis paper was written prior to interacting with any online lessons and one was written after a week of access to one of the specially-designed online lessons. One third of the students participating in the study chose not to access their assigned educational intervention.

The depth of inclusion of key CA elements in their scientific study analysis papers were then used as an indicator of the CA skill level of the students themselves. Trained raters scored the level of CA skill development evident in the student papers using a previously developed CA rubric. A score of 20 out of a maximum of 30 points was considered a basic CA skill competency level. The highest individual paper score in this study was 18.75 which is considered "Developing" on the CA rubric's scale.

This study found no significant differences in student CA skill development between two different treatment groups: an interactive game and an expert video. Qualitative pre- and post-test subject data indicated high student interest in science writing instruction but problems in the timing of the interventions. Logistical and design lessons learned in this study are provided to inform and improve similar future studies.

Chapter 1
Introduction

The students of today need *critical thinking*[1] skills to help solve the complex problems of tomorrow and to effectively participate in policy decisions related to health care, the environment, and other science-related issues (AAAS, 1993; NRC, 1996). The Partnership for 21st Century Skills (p21) advocated that CT instruction be part of the Common Core Curriculum (p21, 2011) underscoring its importance in K-12 education (iNACOL, 2011). While it is easy to recognize the importance of CT, it is difficult to briefly define such a complex concept and experts continue to debate exactly what constitutes CT.

One of the simplest definitions of CT comes from Robert Ennis, a co-author of the highly-regarded *Cornell Critical Thinking Test*, who has consistently defined CT as "reasonable, reflective thinking focused on deciding what to believe or do" (Ennis, 1991 & 2013). In 1990, researchers who developed *The California Critical Thinking Skills Test*, convened a conference of almost fifty CT experts from the fields of philosophy and education. They used the Delphi method to develop a definition and then strove for consensus among these experts. They all agreed that CT is "purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological or contextual considerations upon which that judgment is based…" and "a liberating force in education

[1]1 Definitions of italicized terms appear in the **Definition of Terms** on page 17

and a powerful resource in one's personal and civic life" (Facione, 1990, p. 6). In a more recent paper, Facione elaborated key characteristics of CT as the "habitual intention to be truth-seeking, open-minded, systematic, analytical, inquisitive, confident in reasoning and mature in making judgments" (Facione, Facione & Giancarlo, 2000). Likewise, the Foundation for Critical Thinking has defined CT as "self-directed, self-disciplined, self-monitored, and self-corrective thinking. It presupposes assent to rigorous standards of excellence and mindful command of their use." (Paul & Elder, 2003, p. 2).

These sample definitions illustrate that CT is more than simply an ability to think clearly or logically. In his dissertation, Edward Glaser, one of the creators of the *Watson-Glaser Critical Thinking Scale*, suggests that critical thinking requires three things: (1) an attitude of being disposed to consider in a thoughtful way the problems and subjects that come within the range of one's experiences, (2) knowledge of the methods of logical inquiry and reasoning, and (3) some skill in applying those methods (Glaser, 1947). Similarly, the Foundation for Critical Thinking states that CT "can be seen as having two components: 1) information processing skills that generate beliefs, and 2) the habit, based on intellectual commitment, of using those skills to guide behavior (Scriven & Paul, 1987). This CT disposition comes from the conscious cultivation over time of characteristics such as open-mindedness, fairness, empathy, inquisitiveness, flexibility and respect for other opinions or points of view (Elder & Paul, 2007; Ennis, 1991; Facione et al., 2000). Along with creativity, flexibility, good communication skills and a willingness to be a lifelong learner, CT is considered to be one of the habits of mind (Costa & Kallick, 2000).

Experts agree CT is not an innate skill, but instead it "only appears when students are trained based on specific sub-skills" (Astleitner, 2002, p. 57). These *critical analysis* (CA) skills are hallmarks of CT that are easier to teach and assess than is CT itself (Astleitner, 2002). They include skills such as critical reading and science writing that are the focus of this particular study. Explicitly teaching CA skills can enhance student cognitive skill development (Marin & Halpern, 2011) and provide scaffolding as needed for students asked to critically read, evaluate and summarize their own judgments of primary research literature in writing (Flynn, 2011; Janick-Buckner, 1997; Stav, Nielsen, Hansen-Nygard, & Thorseth, 2010). These skills, along with the disposition for CT, can be modeled for students by experts in a specific field (Lee & Spector, 2012; Jonassen, Carr & Yueh, 1998; Pedersen & Liu, 2001). For example, having experts explain their thinking while reading or writing scholarly papers can provide a model for enhancing these same CA skills when students write their own papers (Smith & Tanner, 2010).

Instructors in content-heavy college science courses are usually reluctant to relinquish live class time for CA skills instruction (Ross, Burgin, Aitchison, & Catterall et al., 2011). The researcher had confirmed this in a pilot study where she had created in-class interventions using classroom response systems (CRS) on these same CA skills to prepare students for this same type of scientific study analysis assignment. The CRS and instructor-led discussion interventions in that study took only ten to fifteen minutes to complete, yet the science instructors in that study asked for future resources to be available to students outside of class time (Adams & Columba, 2014). Therefore, the CA skill interventions in this study will be asynchronous online educational resources.

To measure CT, the aforementioned traditional inventories provide broad measurements in overall thinking patterns. However, they are somewhat costly and may be too general in scope for the kind of targeted instruction (Tsui, 2002; Paul & Elder, 2003; Saxton, Belanger & Becker, 2012) outlined in this proposal. The study proposed in this paper will measure CA skills within students' written analyses of articles from the primary scientific literature, a common college science course assignment (Huerta & McMillan, 2004; Libarkin & Ording, 2012) that proved to be useful for this purpose in the pilot study (Adams & Columba). The CA skills demonstrated within each student paper will therefore be assessed using a CA rubric, previously designed for this type of assignment (see Appendix A) and explained in more detail later in later chapters.

Statement of the Problem

Critical thinking is important for the future, but uncommon in today's workforce. In a recent National Research Council report, over 90% of employers cite critical thinking (CT) and problem solving as "very important" skills in today's workplace. Yet, these same employers found that only 28% of the new college graduates they hired were proficient in these skills (NRC, 2005). Written communication was also deemed "very important" for job success for almost all of these employers, but again, only 16% of the four-year college graduates who recently entered their companies were found to be adequate writers (NRC, 2005).

In addition, our country needs more scientists to solve problems and drive innovation, helping to create new jobs and continue to compete in the global economy (AAAS, 2011; U.S. DCESA, 2011). Yet traditional science curricula have done little to nurture young people's innate curiosity and creativity, two important characteristics of scientists (NSTA, 2011). The scientific process and CT skills needed for developing research skills and *scientific literacy* are often missing as well in today's curricula (Annetta & Shymansky, 2006; Fulton, 2012; Lord & Orkwiszewski, 2006). Students interested in careers in science, technology, engineering or math, aka the STEM fields (NSF, 2004) need more than just research or design skills (NRC, 1996), CT, curiosity and creativity. They also need excellent communication skills to fully understand what has already been discovered and to share the implications of their new discoveries as required (Banko, Grant, Jabot, McCormack & O'Brien, 2013; NSF, 2004). This means that CA skills, such as the abilities to critically read, analyze and summarize primary research literature, are especially important to budding scientists. However, scientific process and communication skills are often neglected in both secondary and post-secondary science education (Banko et al., 2013; Bybee et al., 2006; NSTA, 2011). Instead instruction in these vital CA skills is often replaced by an overemphasis on learning and memorizing scientific content and facts (Baker et al., 2008; Dangel & Wang, 2008; Greenstein, 2013; Khan, 2012).

Recently the science education standards have started shifting toward engaging students in authentic processes of science and problem solving (Banko et al., 2013: NSTA, 2013). This shift is expected to result in deeper understanding and enhanced CT levels in the next generation of scientists as it helps to increase the number of scientifically-literate citizens. The challenge then for science educators is to continue to discover and use effective instructional methods for facilitating the development of CT skills in all students to help prepare them for an unknown future in a rapidly-changing world (NSTA, 2013; p21, 2008).

Nature of the Study

This proposal outlines a two-phase study of the effects of *interactive*, *online learning* on the development of CA skills in college science students. College science students were given a list of five key CA elements, adapted from material developed by the Foundation for Critical Thinking (Paul & Elder, 2003), and asked to address these elements when preparing written analyses of two scientific research articles.

Students could use the list of identified CA skills to guide their reading of the scientific study papers from the primary literature and again when writing their analysis papers, both performance-based tasks. The papers were completed in two phases and students were asked to rate their perception of their CA skills prior to receiving the full assignment instructions. In phase one, the students were assigned to read one of two different articles from the primary research literature and prepare a written analysis of it using only the information included in the written assignment instructions.

Participating students were then randomly assigned to one of two groups to complete the second phase of this study. Each group was granted access to a different type of online instructional intervention, each of which provided information related to the same key CA elements, but with varying levels of content *interactivity*. The first group was given access to a simple educational game that required the students to interact with it in order to proceed through the material. The second group was granted access to a video which was comparatively passive. This video was a montage of clips from interviews with expert biological scientists explaining how they determine the credibility of the scientific papers they must read to stay current in their fields. Students could then apply what they learned in their online educational intervention to the task of writing a second scientific study analysis paper. They were then asked to rate their CA skills again after they had submitted their second paper.

Trained raters reviewed both papers in a blind rating process. Both papers from each student were anonymized and separated from each other so that they were independently included in the totally random order used in the rating process. The raters used the previously mentioned CA rubric, specifically designed for this purpose, to assign a score for each of the five criteria of the rubric met to each of the student papers. The five criteria of the rubric corresponded with the five identified key CA elements. The rater's scores depended on the level of inclusion of these key CA elements within each student paper. The scores from each rater for the individual CA elements were then summed to determine an overall CA score for each paper. The difference between an individual's overall CA score on their first and second paper serves was to serve as an indicator of the level of CA skill development demonstrated by each student. Differences between the individual element scores from each matched pair of papers were also analyzed. Additional quantitative and qualitative data related to CA skill development were collected from each participating student. Further details of this mixed-methods study are outlined in the third chapter of this dissertation.

Research Questions and Hypotheses

This study was designed to measure the main dependent variable of student CA skill level while answering the following research questions (RQs):

1. How do early-year college student perceive their own CA skill levels and do they perceive a need for further development of these skills?
2. Does online student-content interactivity enhance student CA skill development as demonstrated within written analysis papers?
3. Which of five key elements of CA is most affected by variations in online student-content interactivity?

For RQ 1, the researcher hypothesizes that the students will know that they need better CA skills and will be interested in improving them. For RQ 2, the researcher hypothesizes that the gaming intervention will more effectively enhance student CA skill development than will the video, since it will be more interactive which should more fully engage the students (Dunlap, Sobel & Sands, 2007; Kandel, 2006). The researcher further predicts for RQ 3 there will again be significant differences among the treatment group effects on the individual CA elements (Halverson, Siegel & Freyermuth, 2010), as was found in a pilot study completed by the researcher (Adams & Columba, 2014).

Purpose of the Study

The main purpose of this study was to compare the effects of alternative online learning technologies on CT skill development in students taking a first-year genetics course. Results of this study can help determine which types of online interactivity have the greatest impact on student CA skill development. Instructional designers can use the results to inform the design of effective online lessons and help students prepare for successful careers in the STEM fields (NSF, 2004). The end result should improve our collective futures by ensuring we have enough bright scientists, technologists, engineers and mathematicians to develop the new processes and products needed to stimulate innovation and economic growth (AAAS, 2011; OECD, 2008).

Theoretical Basis

The researcher created the online interventions for this study based on constructivist learning theory (Jonassen et al., 1998; Sitzmann, Brown, Kraiger & Kanar, 2009). Constructivists believe students extend their knowledge by making connections between what they have previously learned (also called prior knowledge) and new material (Hannafin, Hannafin & Gabbitas, 2009; Jonassen & Land, 2012). Since students have different levels and kinds of prior knowledge, they will respond differently to new material. A well-designed online constructivist learning activity will reinforce main concepts, test each student's command of related knowledge, and encourage them to move on to more challenging material (Sitzmann et al., 2009). It will also provide enough information to support meaningful learning for students who are not as familiar with the material being presented (Hannafin, Hannafin, Land & Oliver, 1997; Jonassen, Howland, Marra & Crismond, 2008; Khan, 2012).

The planned study interventions were meant to enhance CA skills, such as the abilities to critically read a scientific article and apply reason and logic to its contents. These interventions depicted different ways of thinking about the five key CA elements that the students should include to help create a logical progression within their writing that purposefully builds valid conclusions. (Huerta & McMillan 2004; Libarkin & Ording, 2012). Instructors in content-heavy college science courses are usually reluctant to relinquish live class time for CA skills instruction (Ross, Burgin, Aitchison, & Catterall et al., 2011). Providing online resources to supplement live class instruction has proven to be an effective alternative (Appana, 2008; Annetta & Shymansky, 2006; Fish & Wickersham, 2009), as long as the online learning environment also supports the course learning objectives (Clary & Wandersee, 2012; Poniastowski, 2012). In addition, the fact that online enrollment grew 21% in 2010 while overall college enrollment grew only 12% (Allen & Seaman, 2013), indicates that today's college students are more open to the option of taking online courses as well as or along with classroom based instruction (Bichsel, 2013; Saad et al.; 2013).

These online interventions were also designed with authentic learning tasks designed to engage the student's attention enough to be stored in long-term memory, the first step in learning (Edelson & Reiser, 2006; Lord & Orkwiszewski, 2006). Information is said to have been learned when it is embedded in long-term memory, retained for a long period of time, and recalled, or transferred, when needed in new situations (Sousa, 2011; Tennyson & Rasch, 1988). The experimental interventions were designed to explore the impact of different types of online interactivity on these learning processes which are limited by short-term memory (Baddeley, 2003; Kennedy, 2004; Sousa, 2011). For example, the interventions included variations in content presentation such as narration, presentation, and interactive elements to align with the dual code theory of enhancing learning through associations with engaging images or eliciting emotions

(Kousta, Vigliocco, Del Campo, Vinson & Andrews, 2011; Piavio, 2013). Yet they were also designed to align with cognitive load theory (Paivio, 2013; Sweller & Chandler, 1991) so that they didn't overwhelm the limited capacity of the short-term memory areas of the brain (Cook, 2006). These and other related theories will be discussed in more detail in the literature review which can be found in the second chapter of this proposal.

Assumptions

A major assumption of this study was that short-term improvement in a student's CA skill level enhances that person's CT abilities. In fact, it may be that CA skills need to be continually expanded and practiced over time before they can significantly improve overall CT skills. They would also need to be reinforced to become the kind of habits of mind associated with CT (Costa & Kallick, 2000).

Another assumption was that the science-focused CA rubric is a valid tool for evaluating online learning. This assignment-specific rubric was successfully used and validated using inter-rater reliability data by the researcher in a pilot study of the impact of classroom response systems on CA skill development (in publication). However, the previous validation of this rubric was based on inter-rater reliability alone without testing for intra-rater reliability. Therefore, this study will measure and align both to further validate this instrument. It should also be noted that the proposed *asynchronous* online interventions were quite different from the dynamics of an instructor-led, live CRS session, even though the same rubric was used in this study to score similar artifacts.

Study Scope and Delimitations

The online interventions developed were made available to students taking an introductory, 100-level course in genetics during the fall 2014 semester at a small, private university in the Northeastern United States. Only those students who consented to being part of this CA study were able to access the interventions through the university's *learning management system* (LMS) for a limited amount of time. Students who were enrolled in the particular course but chose not to participate in the proposed study received only written information on the scientific study analysis assignment that included a listing of the five key CA elements to include in their final analysis papers.

The grades that the students received on this assignment within the genetics course were determined partially by their inclusion of these CA elements, but also by other factors such as the timeliness, length, originality and overall quality. The instructor allotted 100 points total (10% of the course grade) to the scientific study analysis assignment. The papers were graded by the researcher who used the same point value of 50 points each for both study participants and non-participants.

There may also be inherent differences between students who take introductory science courses, such as this genetics course, during the fall semester when compared to students who take them in the spring semester. Due to time restraints, data will only be collected during the fall 2014 semester and possible semester effects will not be considered in this study.

Limitations of the Study

The generalizability of this study may be limited by the fact that it will involve undergraduate students from only one institution. However, it will be building on the results of a similar, CA instruction pilot study at a different institution that used the same type of assignment and the same CA rubric for scoring final papers (in publication). The other institution involved in the researcher's pilot study is roughly the same size as the university when graduate student enrollment is included. However, the other institution is a publicly-funded, open access institution offering only 100- and 200-level courses and granting only certificates and associate degree (A.S.).

The university that served as the setting for the CA study discussed in this dissertation is a privately-funded, highly selective university that awards B.A., B.S. and A.B. degrees along with graduate degrees through the doctoral level. The undergraduate students who take 100-level courses at the university where this second study took place are much less diverse in regards to age, academic ability and course load than are the students at the other institution where the pilot study was conducted.

Threats to validity

The researcher randomly assigned the students to intervention groups prior to the start of this study. However, there was no way to ensure that the assigned students accessed the interventions, actually interacted with the various online educational interventions or used any information from them when writing their second scientific study analysis paper.

There was also no way to prevent students from sharing intervention URLs or passwords to allow students from other treatment groups to access more than one intervention. This could have led to the creation of additional independent variables, although this does not seem to have happened in this particular study as discussed in more detail in the fourth chapter of this dissertation which contains the study results.

Analytics can be extracted from LMSs that allow researchers to determine which students accessed each intervention and the total amount of time that the intervention remained open on their computer or other device. However, being open on someone's screen does not necessarily mean they were paying attention to it or even in the room.

Students may also share their paper analyses with each other, as might happen in any type of writing assignment. The possible impact of this will also be tracked by having the students submit their papers online through Lehigh's plagiarism detection filtering program called Turnitin©. This process helped eliminate from the study any papers that were found to be too similar to be considered original work.

Significance of the study

The nature and amount of information in today's complex and rapidly changing world means that CT skills have become "essential skills for the kind of scientific literacy and comprehension" (Gunn, Grigg & Pomahac, 2006, p. 3) needed by all students to meaningfully "participate in and contribute to today's society" (NSTA, 2013). Everyone needs better CA skills to evaluate and process the overflow of information available today and the logarithmic growth projected for the future (Çavdar & Doe, 2012; Costa & Kallick, 2000; Webber et al., 2005).

The Partnership for 21st Century Skills (p21) anticipates that there will be an increased need for CT and CA skills in almost all careers of the future (p21, 2008). Excellent CT skills are required of students planning careers in science, engineering and other STEM fields (NSF, 2004; NSTA, 2013) so that they can stay abreast of relevant research findings, plan new processes or structures, and conduct meaningful research (AAAS, 1993; Ross et al., 2011). This is particularly important because our nation needs more students in the STEM fields to maintain our decades-long economic edge in tomorrow's more competitive global economy (p21, 2008; AAAS, 2011).

Most online learning research to date has been focused on proving that online instruction is as good as live, classroom-based instruction. In addition, educational research studies have mostly studied the effects of the online learning environment on overall course objectives and interactions between the instructors and the learners, without directly measuring the impact on CA skill development resulting from the kind of *hybrid instruction* proposed for this study (Fan & Geelan, 2013; Northrup, 2001; Tallent-Runnels, Thomas, Lan & Cooper, 2006).

Industry trainers and instructional designers are warned to ensure that online course content is interactive (Allen, 2014; Zhang & Zhour, 2003). Yet more research is needed on which aspects of interactivity have the most positive effects on learning in general and CA skill development in particular (Kennedy, 2004; Zdravkovic, 2010). Asynchronous, linear tutorials that ask learners to interact only by advancing presentation slides and taking quizzes have not taken full advantage of the unique affordances of personal computers (Jonassen, 1998). In many cases, they have also created a poor impression of *e-learning* in general among many adult learners (Allen, Dirksen, Quinn & Thalheimer, 2014). This study will attempt to establish some of the key aspects of online student-content interactivity that impact the development of CA skills in college biology students by incorporating best instructional design practices for online learning into the planned educational interventions. The new online interventions from this study that most positively impact student CA skill development will later be made available to interested educators at this university and beyond as an open educational resource (https://www.**oer**commons.org/).

In short, the results of this study will help increase our understanding of instructional design features that truly engage students, enhance CT skills and create positive attitudes toward e-learning (Allen et al., 2014; Zdravkovic, 2010). The study findings will add to our growing understanding of best practices within online learning and training programs at a time when demand for this type of instruction is rising exponentially (Allen & Seaman, 2013; Bichsel, 2013; Boettcher, 2013; Saad, Busteed & Ogisi, 2013).

Definition of Terms

This study involves complex cognitive processes and educational processes that are not easily defined and often still evolving. Working definitions are provided here for some of the most relevant terms.

Asynchronous online learning: Educational content and tools, made available to select learners on the internet for them to access at their own convenience (Allen & Seaman, 2010).

Critical analysis skills: A wide variety of skills that are the hallmarks of CT, including the abilities to question, analyze, evaluate, and reflect on information before making a judgment or taking a personal stand on it (Paul & Elder, 2007).

Critical thinking: One definition is a complex set of cognitive processes that can be recognized in the products of thinking such as written or verbal communication (Moon, 2008). Other definitions can be found in the introduction of this proposal.

e-Learning: Electronic instructional material delivered via the internet. Originally called distance learning, now often used synonymously with online learning (Allen, 2014).

Face-to-face instruction: Traditional, live classroom meetings of instructors and learners in one physical location, sometimes abbreviated as F2F in both the professional and lay literature. (Tallent-Runnels et al., 2006).

Higher order thinking skills: The most complex cognitive processes on *Bloom's Taxonomy of Learning* (Bloom & Krathwohl, 1956). Includes CT used to analyze, evaluate and synthesize information and new ideas (Anderson & Krathwohl, 2001)

Hybrid instruction: Courses that include significant amounts of both online learning and regularly-scheduled *face-to-face* class meeting times. (Means, Bakia & Murphy, 2014).

Interactivity: An attribute of e-learning that stimulates action on the part of the learner, and is meant to enhance engagement with content (Allen et al., 2014; Fan & Geelan, 2013).

Learning management system (LMS): Web-based software platform that supports both *synchronous* and asynchronous learning activities and instruction while tracking learning performance indicators and analytics (Davis, Carmean, & Wagner, 2009).

Massively open online courses (MOOCs): Online courses offered via specially designed *LMS* platforms that support and track thousands of learners in a one course often by using peer review or automated assessments (Fournier, Kop & Durand, 2014: Siemens, 2005).

M-Learning: Use of mobile devices such as iPads and mobile phones to access online learning content or courses (Nedungadi & Raman, 2012).

Online learning: Internet-based instructional material and social interactions deliberately designed for educational purposes. Originally called distance learning, virtual learning or now most commonly e-learning (Means et al., 2014).

PHP: a scripting language used by programmers making dynamic webpages.

Scientific literacy: Controversial term for "capacity to use scientific knowledge, to identify questions and draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity." (OECD, 1998).

Synchronous online learning: Real-time, pre-scheduled gatherings of multiple students via the internet for educational purposes, with or without the instructor including via avatars in virtual environments (Inman, Wright & Hartman, 2010; Tallent-Runnels et al., 2006).

Chapter 2
Review of the Literature

The proposed project involves several inter-related areas of active educational research. The four main topical areas included in this literature review are: (1) critical thinking and analysis skills, (2) *science literacy*, (3) relevant learning theories, and (4) online learning.

Critical thinking and analysis skills

There is a growing consensus in the United States that all students need to develop *higher order thinking skills* to be prepared for the workforce of the future (NRC, 2012 ; NSF, 2004; p21, 2008). Fostering CT and other higher order thinking skills helps prepare all students to meaningfully and thoughtfully participate in looming public policy debates regarding personal health care, genetically modified organisms and other complex issues based on scientific principles and ongoing research (AAAS, 2011; Gunn et al., 2006). Evaluating and judging new scientific information, much of which can be false, misleading or incomplete also requires CT (Astleitner, 2002: Webber et al., 2005). Therefore, CT is considered one of the most important skills to teach (Bybee et al., 2006; Facione, 1990).

The need for CT skills develops over time through interactions with the environment (Piaget, 1985). Higher order thinking skills, including CT, are not innate, but instead need to be learned and practiced (Greene, 2010; Ennis, 2013). These CT skills are best learned through active means such as the processes of analyzing, evaluating, and synthesizing, as outlined in Bloom's Taxonomy of Learning (Bloom & Krathwohl, 1956;

Anderson & Krathwohl, 2001). The development of CT skills has been positively correlated to coursework in history, arts and the humanities as well as the sciences (Halverson et al., 2010; Maeder, 2007).

Fully-developed CT skills are especially important for those pursuing careers in science, technology, engineering and mathematics, known as the STEM fields (NSF, 2004). For example, good CT skills help engineers and scientists evaluate alternative solutions when solving complex problems. For researchers, CT skills are foundational to the processes of conducting experiments, interpreting results in meaningful ways and determining future research directions (AAAS, 2011; NRC, 2012; p21, 2008).

The complex, cognitive processes of CT result from prolonged practice of critical analysis (CA) skills, such as recognizing problem components, identifying relevant issues and evaluating the credibility of scientific information and studies (Halverson et al., 2010; Tsui, 2002: Webber et al. 2005). In addition, CA skills such as critical writing are vital to scientists applying for research funds needed to conduct their experiments, as well as for allowing them to effectively share their experimental findings with others in the scientific community (Coil, Wenderoth, Cunningham & Dirk, 2010; Samuels & Farstrup, 2006). Laurence Greene summarizes the important role of scientific communication best by saying the "most revolutionary research breakthroughs will be practically meaningless, if they are not effectively communicated, through various types of scientific papers, to targeted audiences that can learn, apply and benefit from them" (2010, p. xix).

The American Association for the Advancement of Science (AAAS) lists the "ability to communicate and collaborate with other disciplines" as one of six core competencies in its latest recommendations for science education (AAAS, 2011, p. 33). The National Science Education Standards emphasize the ability to ascertain research credibility and communicate with other scientists (NRC, 1996), while the NRC's latest framework for science education includes CA skills in three of the eight "Scientific and Engineer Practices" to be taught in K-12 classrooms (NRC, 2012).

As scientific advances continue to transform society, it becomes even more important for all students to develop these types of CA skills (Moon, 2008; NSTA, 2011). In this study, two CA skills were emphasized: critical reading and critical science writing.

Critical reading

Students "elicit meaning from a text" (Marschall & Davis, 2012), through critical reading, an important first step to evaluating the credibility of scientific research study articles. Reading the primary literature introduces students to real-world applications, scientific reasoning and the scientific method (Norris, Falk, Federico-Agraso, Jimenez-Aleixandre, Phillips & Yarden, 2009, p. 406). Students also discover the tentative nature of scientific findings, making them better equipped to evaluate findings from primary research that mainstream media later purports to be facts (Halverson et al., 2010; Terry, 2012). Finally, delving into the primary literature can give students an appreciation for the complexity of scientific research which can be "as much a part of scientific inquiry as are observation, measurement, and calculation" (Norris & Phillips, 2008, p. 233).

However, students struggle with the task of reading primary articles, perhaps because the writing style is so strikingly different from what is found in popular press reports or even standard science textbooks (Norris et al., 2009). In fact, many students liken the experience of reading highly technical scientific articles to needing to learn a totally different language (Greenstein, 2013; Janick-Buckner, 1997). Often, they are unsure of what to focus on within a primary research paper and have reported feeling overwhelmed by the challenge of reading them. Explicit instruction on identifying the most important information that matches the key CA skills within a scientific research articles helps students better understand them and write better analyses (Janick-Buckner, 1997; Marin & Halpern, 2011; Marschall & Davis, 2012; Samuels & Farstrup, 2006).

Critical science writing

Excellent writing skills are especially important for procuring the monies needed from potential funders to set up and maintain scientific research laboratories (Greene, 2010). In addition, these critical writing skills are needed by scientists to share their research findings. Most often this is done by publishing both methodology and results of scientific work in peer-reviewed journals with enough detail to replicate the studies, help prove theories and stimulate more scientific advances (Samuels & Farstrup, 2006; Ross et al., 2011).

General college writing courses usually focus on communicating ideas with an emphasis on mechanical skills such as style, coherence, sentence structure, and grammar (Aichison & Lee, 2006; Ellis, 2004). These are important basic skills for all students, but the kind of scholarly writing expected in science is quite different (Greene, 2010; Ross et al., 2011). The main goals of science writing are to clearly articulate new ways of thinking of new experimental findings or past research studies to build logical arguments that support experimental theories and suggest promising directions for new research studies (Coil et al., 2010; Ondrusek, 2012; Ross et al., 2011). In other words, science writing is not so much about recording what has already happened as it is about generating new research, new insights, and ultimately new understandings about our world (Aichison & Lee, 2006; Meyer and Land, 2005; Tardy, 2005).

Educational scaffolding that enhances critical writing skill development in students helps enhance critical thinking and science literacy (Corradi, 2012; Ediger, 2012; Halverson et al., 2010). To generate knowledge from writing, would-be scientists need to remain open-minded as they evaluate sources of credible scientific information (Ellis, 2004; Greenstein, 2013). They need to continuously evaluate their own thinking to ensure a logical progression towards understanding the issues before taking any personal stands or investing in new directions for their research (Ross et al., 2011). Developing science-writing skills takes practice and requires instructional resources that support the development of student writing skills along a continuum from clarity to new insights (Ediger, 2012; Ondrusek, 2012; Samuels & Farstrup, 2006).

The writing-to-learn (WTL) movement appears to be an effective way to improve academic writing skills in K-16 students (Balgopal & Wallace, 2013; Fry & Villagomez, 2012; Nevid, Pastva & McClelland, 2012; Quitadamo & Kurtz, 2007). Instead of focusing on the finished product (e.g. a paper), WTL focuses on organizing information, articulating original ideas and building supportive learning communities. This popular approach closely mimics the writing processes of scientists (Ke, & Hoadley, 2009; Ross et al., 2011). In fact, the emphasis of WTL is on the powerful, cognitive processes of paraphrasing, which expects students to convert difficult material into their own words.

This requires them not only to know the material, but also to think deeply about it (Greenstein, 2013; Marschall & Davis, 2012). Some researchers feel that WTL is an underutilized tool at all educational levels, including post-secondary education (Balgopal & Wallace, 2013).

Another effective approach to enhancing critical writing skills is Calibrated Peer Review,™ also called CPR. This web-based software program contains a repository of learning activities designed to enhance CT and critical writing skills of college students by simulating the peer review process of academic publishing (Fosmire, 2010; Gunersel, Simpson, Aufderheide & Wang, 2008). With CPR, students evaluate and recommend the work of their classmates while the software automatically tracks and grades their efforts. Developed by Orville Chapman, a chemistry professor at UCLA, CPR (cpr.molsci.ucla.edu) is available for purchase on an institutional level. The CPR model has been successfully used at over 700 educational institutions, but it is relatively expensive. In addition, some users feel CPR needs more high-quality critical writing activities in its learning objects library (Reynolds & Moskovitz, 2008).

Summary of critical thinking and analysis tools

All students need to have the ability to think critically when evaluating the ever-changing world around them, but CT is especially important for those planning careers in the STEM fields. Explicit instruction in CA skills such as critical reading and critical science writing can greatly enhance development of overall CT skills.

Scientific literacy

Science faculty members agree that CA skills are important, yet they feel compelled to focus on building student knowledge and understanding of scientific content. Most often this is accomplished through lectures and readings that require rote memorization of facts (Fraser, Timan, Miller, Dowd, Tucker & Mazur, 2014; Greenstein, 2013), resulting in an overemphasis on content knowledge in science classrooms (Khan, 2012; Wiggins & McTighe, 2005). Focusing on lower order thinking skills, such as remembering and understanding, means less attention given to developing higher order thinking skills, such as CT. Yet, these skills are the ones most needed for understanding the dynamic and complex applications of science to real-world problems and to make new discoveries (Lynd-Balta, 2006; Watkins & Mazur, 2013). Everyone, regardless of future career aspirations, needs basic CA skills to understand scientific methods enough to productively participate in debates surrounding everyday STEM-related issues such as the global warming and energy production (Banko et al., 2013; Fraser et al., 2014).

Teaching too many facts can also give students the impression that science is merely a body of static knowledge (Baker et al., 2008) at a time when “science (and technology) are evolving at a faster and faster pace and content is prone to becoming obsolete“ (Holbrook & Rannikmae, 2009, p.277). Science education is coming under pressure to change as support grows for more instruction in the processes and contexts of science (NRC, 2012; p21, 2008; Woodin, Smith & Allen, 2009). Reform seems especially urgent now since our nation is in need of more workers in the STEM fields to drive innovation, grow our economy, create future jobs and help the U.S. remain globally competitive (NSF, 2004; US Department of Commerce, 2011).

Science education also needs to effectively engage all students to ensure diversity of thought and input into designing our collective future. When too much time is devoted to the presentation of facts, correspondingly less instructional time is left for fostering student CA skills through making observations, designing experiments and other scientific activities that require curiosity and creativity (AAAS, 2011, NRC, 1996). In fact, many budding scientists often get frustrated in content-heavy science courses and end up choosing different fields that they perceive to better value their native creativity and inborn curiosity (Greenstein, 2013; Lynd-Balta, 2006; Ross et al., 2011). Thus an over-emphasis of science content over context may actually be driving away many potential scientists with unique, innovative perspectives.

When science teachers do incorporate problem solving, they traditionally show the students how to calculate the correct solution and then assign similar problems for the students to practice solving as homework. However, most students focus on memorizing the mechanics of the demonstrated solutions without really understanding *why* the particular solution is used or the real-world applications of this type of problem (Watkins & Mazur, 2013). Students usually overemphasize surface features of complex problems and miss recognizing the underlying patterns and concepts used by expert scientists, such as their professors (Jackson, Draugalis, Slack & Zachry, 2002; Smith & Tanner, 2010). Thus, students are left with only “loosely organized, ill-defined bits and pieces of knowledge that are dependent upon the specific circumstances in question” (Huffman &

Heller, 1995, p. 141). Instructional techniques that immerse students in scientific content by requiring them to explain their thinking to others, seem to produce deeper understanding of problem-solving and any underlying scientific concepts (Baker et al., 2008; Fraser et al., 2014; Greenstein, 2013).

Many of today's discovery science learning models are based on Piaget's cognitive and developmental theories and education strategies (Piaget, 1985). This includes the 5-E Learning Cycle which consists of five iterative steps: (a) engage, (b) explore, (c) explain, (d) elaborate and (e) evaluate (Bybee et al., 2006). The 5-E Learning Cycle is used extensively in science education because it invites students to participate in a scaffolded, modified version of the scientific method itself. For example, the first step in the scientific method, observation, corresponds with noticing something new or unexpected that *engages* the student's attention, the first step of the 5-E model. The last step of the 5-E model, *evaluation,* is actually part of all of the other steps (Bybee et al.,2006) allowing the students and the teacher to monitor learning throughout the cycle before starting another iteration. This widely used model has proven itself so effective in enhancing student knowledge and continued interest in science (Bybee et al., 2006) that the newly adopted *Next Generation Science Standards* highly endorses the continued use of it in K-12 science education (Achieve, 2013; NSTA, 2013).

Performance on international science and mathematics examinations by U.S. students has also been steadily declining over the last couple of decades. The U.S. average scores on these key assessments of science education are now below the average scores of all developed countries participating in the exam (OECD, 2008). American policy-makers and business leaders have been clamoring for cutting-edge science education to prepare students for unknown future challenges and jobs that may not yet exist (Collins & Halverson, 2009; Woodin et al., 2009). Some business leaders and companies have joined forces with leading STEM educators to form the Partnership for 21st Century Skills (p21), with CT development as a major component of innovation in its *21st Century Skill Framework* (p21, 2008). This partnership's follow-up document, *Common Core Toolkit,* advocates classroom activities that allow students to gain experience in problem-solving, using technology and collaborating to better mimic the workplace of tomorrow (p21, 2011). In response to p21's recommendations, the National Science Teacher Association (NSTA) endorsed these same 21st Century Skills (NSTA, 2011, p21, 2008) and incorporated them into their newly released *K-12 Next Generation Science Education Standards*, which have also been endorsed by several key states (NSTA, 2013).

Science educators agree scientific literacy is important for all students, most of whom will never go on to become research scientists. Although there is no standard definition of this term, there appears to be two main aspects of scientific literacy. The first is mastery of scientific content (Bybee, 1997) and the second is a thorough understanding of the tentative, evolving nature of scientific research (Clary & Wandersee, 2012; OECD, 1998). This second aspect is needed by all students to make informed decisions, although most experts maintain that both aspects are very important for developing a scientifically literate citizenry (AAAS, 1989; Holbrook & Rannikmae, 2009; NRC, 1996). Norris and Phillips (2003) developed a similar context-oriented framework containing a fundamental and derived sense of science literacy. The fundamental sense is related to the importance of scientific reading, writing and communication in general contending that these skills are intertwined with science meaning and learning. The broader, derived sense of science literacy is related to science knowledge that "individuals need to function successfully in society" (Norris and Phillips, 2003, p. 229). This has led to a debate about which aspect of science literacy is most needed to ensure correct and appropriate future applications of scientific knowledge when needed by individual members of society (Albitz, 2007). After reviewing over fifty relevant papers, Fan and Geelan (2013) emphasized that teaching science processing skills over science content is more likely to lead to this overall result, although some science content knowledge is also important (Fan & Geelan, 2013).

Examination of the thinking processes of expert scientists used in interviews about their work and in their scholarly writing demonstrates the dynamic nature of scientific knowledge itself and seems to favor a deeper sense of understanding of the processes of science (Halverson et al., 2010; Smith & Tanner, 2010). This type of understanding of science seems especially urgent in the face of current, scientifically-based public policy debates on critical issues such as genetically-modified foods, appropriate energy sources and global warming (Albitz, 2007; Poniastowski, 2012). This movement to emphasize the underlying thinking processes of science for all is a key component of the recently-released *Next Generation Science Standards* (NSTA, 2013).

In an attempt to help college students and educators interested in improving the science literacy skills of the general population, the Foundation for Critical Thinking created "*A Checklist for Scientific Reasoning*" (Paul & Elder, 2003, p. 4-5) outlining the following foundations of scientific thinking and understanding:

- All scientific reasoning has a PURPOSE.
- All scientific reasoning is an attempt to figure something out, to settle some scientific QUESTION, or solve some scientific PROBLEM.
- All scientific reasoning is based on ASSUMPTIONS.
- All scientific reasoning is done from some POINT OF VIEW.
- All scientific reasoning is based on scientific DATA, INFORMATION and EVIDENCE.
- All scientific reasoning is expressed through, and shaped by, scientific CONCEPTS and THEORIES.
- All scientific reasoning contains INFERENCES or INTERPRETATIONS by which we draw scientific CONCLUSIONS and give meaning to scientific data.
- All scientific reasoning leads somewhere or has IMPLICATIONS and CONSEQUENCES.

Students planning to pursue careers in the sciences will find that incorporating the reasoning skills listed above helps them to evaluate scientific information from any source including professional, peer-reviewed articles (Mauldin, 2011). Many science teachers attempt to enhance scientific literacy in the classroom by asking all of their students to use a scaffolded version of the expert reasoning skills listed above when judging the validity of scientific findings reported in the media (Hoskins, 2010; Maudlin, 2011; Terry, 2012). Students can also enhance their own CA skills by applying these expert reasoning skills to the task of critically reading, analyzing and summarizing new scientific information, most often provided to all by the popular media (Flynn, 2011; Stav et al., 2010).

However, it can be difficult to fully evaluate scientific information within lay literature because clear citations or links to original, foundational research studies are rarely, if ever provided (Halverson et al., 2010; Hoskins, 2010; Terry, 2012). Students also get confused when media reports include contradictory results from different scientific studies. Teacher Sally Hoskins (2010) suggests K-12 students be taught to determine strengths and weaknesses of scientific reasoning by considering the following topics while reading scientific research studies: (a) correlation vs. causation, (b) variables & controls, (c) soundness of data, (d) scientific skepticism, (e) nature of science, (f) experimental design, (g) science and society. Identifying these factors is a pre-requisite for most students' abilities to decide for themselves the soundness of the conclusions drawn and the validity of any resulting scientific claim (Greene, 2010).

The proposed reforms in science education will require new professional development programs for K-16 educators. These new programs need to emphasize scientific *technological, pedagogical content knowledge* or *TPACK* needed to model and more effectively teach these 21st century skills (Mishra & Koehler, 2006). Improving teacher technology skills provides them with the confidence needed to more fully utilize the unique affordances of the internet to enhance 21st Century skills, such as collaboration and self-directed learning, within their students (Ward, 2010).

Summary of scientific literacy

When science educators do not challenge their students to practice science processing skills, they jeopardize the development of the kind of scientific literacy, now needed by all students (AAAS, 1993; Lynd-Balta, 2006; Poniastowski, 2012) and decrease the likelihood that a wide spectrum of students will be attracted to STEM careers (Greenstein, 2013; Lynd-Balta, 2006; Ross et al., 2011).

Relevant Learning Theories

Educators agree that online instruction and educational resources should be based on recognized educational theory (Merrill & Drake, 1996). However, "rigid adherence to a particular perspective limits our capacity to distinguish proven [learning systems] from trendy approaches" (Hannafin et al., 1997, p. 114). Therefore, the researcher has chosen to review a combination of learning theories that will be used to ground and inform the design of the learning interventions outlined in this proposal. The foundational one will be constructivism.

Constructivism

The theory of constructivism contends that learners build their understanding of new ideas based on what they have already learned and on what they discover through new interactions with the world around them or within learning environments. Piaget was one of the first educational theorists to articulate the idea that learners construct an understanding of the world based on their experiences within it. He proposed that an uncomfortable state of cognitive dissonance is created when learners discover new facts that do not fit with their preconceived notions, which he called schema. He theorized that for learners to return to equilibrium, and become comfortable again, they must modify their pre-existing schema via the processes of assimilation, accommodation or reordering of knowledge within their mind so that they can incorporate the new-found knowledge and reconcile the cognitive dissonance and return to equilibrium (Piaget, 1985).

David Jonassen, a more recent leader in the field of constructivist learning theory, called computers "mindtools" that can help construct knowledge (Jonassen et al., 1998). He was the first to note that hyperlinks in online learning environments mirror the way that the human brain connects new ideas to prior knowledge in order to build a web of new understandings (Jonassen et al., 1998). He also wrote extensively on how education can benefit by more fully utilizing the internet and other technological advances to create rich, scaffolded learning environments in which students can explore concepts and create their own understandings (Jonassen et al., 1998; Jonassen et al., 2008). Jonassen has proposed that these constructivist learning environments will most effectively support learning when they are active, intentional, complex, contextualized, reflective, conversational and collaborative (Jonassen & Land, 2012).

Cognitive learning theory

This learning theory proposes that the storage of new knowledge in long-term memory is the basis of learning (Sousa, 2011; Tennyson & Rasch, 1988). Advances in neuroscience have recently validated some of the premises of this learning theory which impacts teaching and learning. First, anything that sufficiently engages student attention will be passed along neuronal pathways to the immediate short-term memory area of the brain, while all irrelevant or mundane information will be discarded (Sousa, 2011).

Therefore, instruction must capture the student's attention by being relevant or important enough to the student for it to be processed (Kandel, 2006). Once new information has been sent to immediate short-term memory, inter-neurons actively compare the incoming data to past experiences for about thirty seconds before determining if the new data will be passed along or dropped from the system (Kandel, 2006). Repetition of key concepts by the instructor or mental rehearsal of new information by the student can help retain the information in working memory. Focused attention is then needed to associate this new material with prior knowledge and the average attention span for college students appears to be between 10-20 minutes long (Brunce, Flens & Neiles, 2010). Creating short online learning objects takes full advantage of student abilities to focus on learning new material (Brunce et al., 2010; Sousa, 2011).

However, unless it has meaning to the student and makes sense to them by fitting in with their view of the world (also called their *cognitive belief system*), it will start to fade away and be forgotten soon after it stops being rehearsed (Sousa, 2011). Educators can help students remember information at this stage by using real-world examples, case studies and problems that align with what students find meaningful or already know. This will greatly enhance the likelihood that the desired content will be learned and retained (Allen, 2014).

Information related to personal survival or that ilicit strong emotions will also automatically have more meaning to the learner and thus be preferentially retained (Sousa, 2011). If the scenarios or examples used by the instructor elicit an emotional response or are deemed important for survival by the students, they have an even greater chance of being stored in the students' long-term memories and therefore truly learned (Sousa, 2011).

Dual code theory

Neuroscience research suggests that working memory has two subsystems: a visuo-spatial sketchpad that handles incoming visual cues; and an independent phonological loop that manipulates incoming sounds and speech-based information (Baddeley, 2003; Paivio, 1969). Both are under the control of a central executive subsystem that monitors input from these subsystems and makes connections between the information being relayed and long-term memories (Piavio, 1969; Sousa, 2011). Items with additional association triggers, such as images and sounds, therefore have an increased probability of being encoded, stored in long-term memory and later recalled along with these triggers (Baddeley, 2003). Still others feel that emotions have such a strong imprinting effect on memory that there may actually be additional channels controlling access to working memory (Kousta et al., 2011; Paivio, 2013). While this issue is still under debate, instructional designers continue to apply this theory when developing learning objects by combining words and images, along with sound, to simultaneously stimulate more than one sense and powerfully enhance recall of instructional content (Paivio, 2013; Moreno & Mayer, 2002).

Cognitive load theory.

Working memory is the limiting factor in the cognitive processes of knowledge acquisition by the brain (Baddeley, 2003; Sousa, 2011). Instructional designers must be careful when adding additional educational features such as images, graphics or sound, to not overstimulate the senses and make it more difficult for the intended learners to process the materials (Cook, 2006; Sweller & Chandler, 1991). If the cognitive load of the educational materials overwhelms the working memory's limited capacity to process them, students will not be able to learn the material (Mayer & Moreno, 2003). For example, if the words and images within a lesson are placed too far apart, the learner will need to hold one word or image in working memory while looking for a corresponding match. Better spatial arrangement of the instructional content can prevent this situation, decrease the cognitive load and therefore enhance learning (Sweller & Chandler, 1991; Mayer & Moreno, 2003).

Summary of learning theories

Various learning theories can be effectively combined to create a variety of learning methods and outcomes to help develop CT in students (Fish & Wickersham, 2009; Moore, 1989). The researcher plans to incorporate several of the guidelines listed above in the experimental interventions designed for CT instruction in this study as outlined in the third chapter of this proposal. The resulting learning objects can then be used for specific educational purposes in a multitude of settings from individual study to *m-learning* and beyond (Nedungadi & Raman, 2012; Northrup & Rasmussen, 2000).

Online Learning

In higher education, online learning has been growing faster than overall enrollment in the last decade and is projected to continue growing exponentially in the foreseeable future (Allen & Seaman, 2013; Bichsel, 2013; Saad et al.; 2013). In the eighth annual online learning report from the Sloan Consortium, it was reported that in

2010 online enrollment grew 21% when overall college enrollment grew only 1.2% that year (Allen & Seaman, 2013). In a recent telephone survey of over 1,000 college presidents, conducted by the Pew Research Center, 89% reported that their institutions now offer online courses, although only 69% of privately-held four-year colleges reported doing so (Parker, Lenhart & Moore, 2011). Over 6.7 million students, about one third of all college students in the country, enrolled in at least one online course in 2010, which is over five times the total number of students who reported doing so in 2002 (Allen & Seaman, 2013).

Most schools (79%) not yet offering online courses at the time of the survey were planning to do so in the near future (Bichsel, 2013). In addition, half of the college presidents predicted that by 2021 most of their students will be taking online courses (Parker et al., 2011). Most of the growth reported in online course enrollment was generated by institutions with established online learning programs rather than from more traditional institutions starting new online learning programs (Allen & Seaman, 2010).

Online learning increases access to higher education by being more affordable than traditional classroom options and offering expanded educational topics. Asynchronous courses provide learning opportunities almost anytime and anywhere with internet access which is attractive to students who travel a lot, are home-bound, or have unusual schedules (Allen & Seaman, 2013; Appana, 2008; Bichsel, 2013; Collins & Halverson, 2009; Parker et al., 2011; Saad et al., 2013).

However, many college instructors are comfortable with the status quo of traditional *face-to-face instruction* and hesitant to adopt and embrace any new technological advances (Rogers, 2003). This is especially true for online learning since it takes a lot of time to develop an online course shell and college instructors usually have multiple demands on their time (Amiel & Orey, 2007). In fact, the "major limitation to developing online courses is the experience and knowledge of the instructors who have different levels of creativity and technical knowledge." (Appana, 2008, p. 14)

Much of the online learning literature is focused on the differences between social interactions (student-to-instructor or student-to-student) in online learning environments vs. live classrooms (Tallent-Runnels et al., 2006; Ward, 2010). Synchronous online activities provide the closest online approximation of traditional face-to-face classes (Moore, 1989; Tallent-Runnels et al., 2006), while affording flexibility in location, but not in class meeting times. Students and instructors can be halfway around the world from each other, and still participate in a synchronous, interactive educational activity such as a debate, a brainstorming session, a real-time presentation or a meeting about a group project (Anderson, 2004; Collins & Halverson, 2009). However, because of today's busy personal schedules, many online courses have mostly asynchronous activities with only occasional synchronous sessions (Allen & Seaman, 2013; Tallent-Runnels et al., 2006). Asynchronous interactions and discussion forums provide students and instructors with ample time to reflect on whatever was posted and craft researched, thoughtful responses (Greenlaw & DeLoach, 2003: Lapadat, 2006). This may be why, in the relatively short history of online learning, asynchronous discussion forums have proven to be one of the most effective ways of enhancing meaningful online learning (Greenlaw & DeLoach, 2003; Lapadat, 2006; Perkins & Murphy, 2006).

Today, a growing number of educators are eager to follow the lead of their more adventurous colleagues and explore the instructional potential of online learning (Rogers, 2003; Tallent-Runnels et al., 2006). The widespread availability of the internet in classrooms from K-16 has resulted in hundreds of comparison studies over the last few decades. When studies are restricted to only one independent variable by using online learning activities that are similar to classroom resources, as suggested by Clarke (1994), the results of online learning are comparable to those of traditional instruction (Price, Richardson & Jelfs, 2007; Smith & Palm, 2007; Wang & Woo, 2007). The bottom line is that well-designed and monitored online learning has proven to be as effective as traditional, live classroom instruction over a broad spectrum of subjects (Fournier, Kop & Durand, 2014; Means et. al, 2014).

Yet, both the general public and many employers perceive online education as less rigorous and therefore of lower quality than traditional, classroom instruction (Parker et al., 2011; Saad et al., 2013). Even though instructional designers have expanded personalized online instruction, most Americans also think that traditional education is better suited for meeting individual educational needs. These impressions may be due in part to the fact that people understand best what they themselves have experienced (Khan, 2012). This negative public impression may also be influenced by the content-heavy, linear presentations that have been all too commonplace in the field of what is now called e-learning, especially in workplace training programs (Allen et al., 2014; Parker et al., 2011).

Another factor driving public perceptions that online courses are easier and therefore, not of as high a quality as face-to-face instruction, may be the fact that many online assessments and course assignments have been largely unproctored. Online assessments often rely instead on an honor system to ensure students do their own work. This essentially turns every online examination into an open exam in which students can easily look up needed information (Harmon, Lambrinos & Buffolino, 2010; Milliron & Sandoe, 2008; Watson & Sottile, 2010). Evaluating knowledge gains through projects and portfolios that require students to create their own answers can help authenticate learning (WCET, 2009). However, all online instructors, including those leading large groups of students in *massively open online courses* or *MOOCs* must take the necessary steps to minimize students' abilities to cheat or there will be continuing mistrusting in the legitimacy of online student learning assessments (Fournier et al., 2014; Milliron & Sandoe, 2008; Watson & Sottile, 2010).

If leaders in the field want online learning to have equal status with campus-based programs in the public's eye, they need to demonstrate higher standards for testing, grading, and instruction, according to a 2013 Gallup survey report. More rigorous assessment standards should lead to greater employer acceptance of online educational qualifications (Parker et al., 2011, p. 3). Once these changes are in place, greater public appreciation for the benefits of online learning will likely follow (Parker et al., 2011). Better instructional design is key to improving online education while maximizing the full educational capabilities of the internet (Allen et al., 2014; Ward, 2010).

Instructional design for online science learning

Like any other form of instruction, online learning is most effective when it is planned to meet the specific needs of the learners and offers opportunities for individualized instruction (Means et al., 2014; Gredler, 2009; Snowman & Biehler, 2006). Planning implies a goal, which in education usually relates to what the learner should know or be able to do after they have participated in a course, learning activity, or learning environment (Allen et al., 2010; Gredler, 2009; Wiggins & McTighe, 2005). Once a learning goal is established, the instructional designer then needs to decide how best to measure or assess the learner's performance and overall progress towards that goal (Costa & Kallick, 2000). Only when an instructional designer knows what is expected and how it will be measured, can he/she decide what kind of instructional tools or online learning environments are needed (Allen & Sites, 2012; Jonassen et al., 2008; Wenglinsky, 2005).

The bulk of the research in online learning to date has proven it to be comparable and in some cases better than traditional, classroom-based instruction (Khan, 2012; Means et al., 2014; Siemens, 2005). Yet, few peer-reviewed studies have compared learning results achieved from variations in different types of online learning (Twigg, 2001; Ward, 2010). The expected potential of online learning to help individual students learn in ways that transcend face-to-face learning experiences has yet to be fully explored (Reiser, 1994; Ward, 2010). Some of the unique affordances of online learning include the abilities to: (a) customize learning for individual students, (b) provide enhanced learner control which may affect self-efficacy and encourage more interactivity, (c) provide just-in-time learning that is reinforced by its immediate use while the student is still highly receptive to the information, and (d) provide individualized scaffolding as needed (Fisher, Wasserman, & Orvis, 2010; Twigg, 2001).

The relative anonymity of online learning environments seems to encourage many students to more deeply share their thoughts about writing (Çavdar & Doe, 2012; Perkins & Murphy, 2006). Providing critical reading and writing instruction through online learning allows students to review instructional materials and scaffolding on their own schedule without taking any time away from traditional, classroom-based science content instruction. Other researchers have reported similar successes with web-based approaches designed by educators for their own use in individual classes or departments (Kalman, 2011; Lapadat, 2006; Poniatwoski, 2012).

However, not everyone designing educational materials is trained as an educator. Also, since there is no general oversight in place for online learning, anything can be posted to the internet, including advertisements presented as educational information. The result has been a seemingly overwhelming amount of resources that can be biased, misleading or simply wrong, mixed in with credible educational resources of varying levels of quality (Astleitner, 2002). Merrill and Drake were the first to call for online resources to be developed using the design principles of already proven learning theories (1996). At that time, computers were just entering K-12 classrooms in large numbers and they were being used primarily for computer skill drills. Some educators were expressing concern that the vast potential of computers and the internet for education was not being fully utilized (Kozma, 1994). Still other educators maintained that computers were just one more expensive educational technology experiment that was doomed to fail (Clarke, 1994) just as other technology and media trends (e.g. filmstrips or television in the classroom) had in the past (Reiser, 2001). The divisive "media vs. method" debate that ensued was not as much about educational tools as it was about the effectiveness of constructivism versus more traditional instructivism theory (Reiser, 1994).

In the wake of this debate, a group of instructional designers developed the following guidelines for grounded-learning systems design that would "inform online instruction but not argue for the inherent superiority of one theoretical position or methodology over another, but for articulation of and alignment among the underlying principles that define them" (Hannafin et al., 1997, p. 103). They stated that effective online learning must: (1) be based in a defensible theoretical framework, (2) use methods that have been proven through sound educational research to support the chosen theoretical framework, (3) be generalizable so that they can be easily adapted or adopted by other designers, and (4) be validated iteratively through successive implementations. However, much of the early computer-based educational resources were developed by people who were knowledgeable about technology and programming languages needed to originally create online learning objects, but they were not educators and they were not familiar with educational theories (Merrill & Drake, 1996). Today many web-authoring tools do not require knowledge of computer programming and anyone who takes the time to learn these new applications can create online learning resources relatively easily, with or without the supporting educational design theories and research-based best practices in educational design (Allen et al., 2014).

Interactivity

Student engagement with the material being learned is an important aspect of the underlying learning theories previously discussed. Little, if any, learning can occur if the student does not engage and interact with the material (Dewey, 1938).

There are three broad categories of effective online interactions: instructor-to-student, student-to-student, and student-to-content (Moore, 1989). These interactions help engage students, thereby enhancing learning (Kennedy, 2004; Sousa, 2011). Most of the online learning research to date has focused on enhancing effective social interactions and therefore, not as much scholarly research is available on student-to-content interactions, and their effects on learning (Dunlap et al., 2007; Murray, Pérez, Geist & Hedrick, 2013; Northrup, 2001). However, this third type of interaction is important because if a student simply gathers new knowledge without ever interacting with it via reading or reflection, then the new knowledge will never be learned by that student (Moore, 1989; Murray et al., 2013).

Specific types of content-specific interactions were first defined by Stouppé (1998) and then expanded upon by Garrison, Anderson and Archer's Community of Inquiry framework (1999, 2001) to create a list of interactions that fell into two broad categories. The first category deals mostly with mechanistic interactions that allow learners to access material through links, navigate within the material, change its order or manipulate it for better viewing. The second category of student-content interaction types are more directly tied to learning by challenging students to engage in problems, make connections between concepts, apply new ideas, test solutions and learn content through simulations, games and other more active functions (Garrison, Anderson & Archer, 2001).

Later Dunlap, Sobel and Sands (2007) added reflective inquiry and metacognitive interactions to the growing list of interactive e-learning strategies. They also modified a version of Bloom's Taxonomy of Learning (Anderson & Krathwohl, 2001) to create a Student –to-Content Interaction Strategies Taxonomy. They suggested that the most effective e-learning objects and environments use a variety of different interactions (Dunlap et al., 2007). However, it should be noted that even the most interactive, best-designed e-learning resource can improve student learning only if it is accessed and used by students (Khan, 2012; Means et al., 2014). There is some evidence to suggest that students are more likely to access what they consider to be supplemental online instruction when they perceive a strong direct correlation between the time they spend interacting with an online resource and their course grades (Means et al., 2014; Murray et al., 2013). Once a student decides to use an e-learning resource, they need to be challenged to work out problems or concentrate on scenarios that lead to deep learning (Dunlop et al., 2007, Khan, 2012; Sousa, 2011).

Best practices in e-learning development

The best online learning environments provide student-centered, authentic problems or situations that are relevant, engaging and meaningful to the intended learners (Allen et al., 2014; Collins & Halverson, 2009; Keller, 2008). They also provide opportunities for personalized learning and performance enhancement through branched or level-graded opportunities. Finally, they help learners engage with the instructor and other learners as well as with course content (Collins & Halverson, 2009; Moore, 1989).

In the burgeoning world of corporate training and development, high expectations necessitated more structured instructional design models (Jonassen, 2008). Too often the results of the most commonly-used design models have been disappointing, linear e-learning modules that are more like online textbooks (Allen & Sites, 2012; Allen, 2014; Jonassen, 2008). Recently, a group of highly-respected instructional design leaders have focused attention on this issue by releasing a manifesto for change (Allen et al. 2014). They feel far too many e-learning programs focus on knowledge delivery and use only didactic feedback, limiting learning to the lower-order levels of Bloom's taxonomy (Bloom & Krathwohl, 1956). They have dubbed these often ineffective programs "typical e-learning" and are calling for all in the field to instead create only "Serious E-Learning" (Allen et al., 2014). Those who take on this challenge, pledge to design e-learning only when necessary and to use an iterative design process, based on accepted learning theories, that meets the needs and expectations of all stakeholders. Online learning designers who make this pledge further agree to create e-learning programs that incorporate realistic scenarios with real-life consequences, tied to desired performance goals. These leaders of e-learning maintain that the most effective online learning environments encourage learning by making mistakes in a supported, authentic environment, with ample examples and counter-examples reinforced by feedback that corrects common misperceptions (Allen et al., 2014; Northrup & Rasmussen, 2000). They also suggest inserting navigational options to allow learners to skim through material already mastered while providing sufficient guidance for those who need it. In short, these expert instructional designers are advocating for more effective e-learning that respects the learner's time by allowing them to rehearse actual work-force tasks, build confidence in newly-practiced skills and therefore enhance long-term job performance (Allen et al., 2014).

Summary of online learning

Designing e-learning that focuses on content relevant to the learners enhances the likelihood that they will truly engage with the material (Allen, 2014; Anderson, 2004). Students must consciously interact with the information to activate cognitive processes that store new information in long-term memory, and therefore learning (Coates, 2007; Sousa, 2011). It follows, then, that instructional methods with greater levels of student – to-content interactivity should lead to enhanced levels of student CA skill development.

Chapter 3
Methodology

The purpose of this mixed-methods educational research study was to determine the impact of different variations of online CA instruction on the development of CA skills in college science students. Different types of online instruction about five key CA skills were developed specifically for this project and presented through an LMS outside of the customized version of Moodle used by the university. The subjects were all early-year science students enrolled in a traditional classroom section of an introductory (100-level) genetics course at a small, private university in the Northeastern United States.

Research Questions

Again, the study was designed to measure the main dependent variable of student CA skill level while answering the following research questions (RQs):

1. How do early-year college student perceive their own CA skill levels and do they perceive a need for further development of these skills?
2. Does online student-content interactivity enhance student CA skill development as demonstrated within written analysis papers?
3. Which of five key elements of CA is most affected by variations in online student-content interactivity?

The answers to the first question were gleaned from the student's responses to the questions in the study pre-test and post-test as well as through their written analysis papers. The CA scores assigned by the raters to the student papers were used to determine the answer to the last two research questions.

Design of the CA assignment

Complex, cognitive processes like CT, are best measured using a performance-based assessment (Jackson et al., 2002; Saxton et al., 2012) that requires students to produce their own answers, rather than simply recognizing correct answers from a list of possibilities, as happens with typical multiple-choice exams (Ennis, 1991). In this study, the researcher measured the levels of CA skills demonstrated by students within their written analyses of scientific research papers, a common assignment in introductory science courses (Huerta & McMillan, 2004; Libarkin & Ording, 2012). The two primary research articles to be analyzed were chosen by the course lecturer and contained content related to upcoming lectures, rather than on concepts that had already been covered prior to the start of the writing assignment (Saxton et al., 2012). To enhance fidelity, all of the students analyzed the same two primary scientific study papers.

Paper A was a highly-cited comparison study of the DNA methylation patterns of pro-nuclei in zygotes of several different species including humans (Fulka, Mrazek, Tepla & Fulka, 2004). Paper B was a more recent study of a gene suspected of playing a major role in the development of Autism Spectrum Disorders in humans (Oksenberg, Stevison, Wall and Ahituv, 2013). These two papers were similar in that they were both from reputable journals, they were of approximately the same length (five pages of text) and they were based on extensive literature reviews with adequate numbers of credible references. Both primary research articles contained numerous figures and photographs although there were almost twice as many of these in the second article (Paper B by Oksenberg et al., 2013). Both articles required similar levels of technical understanding, which was challenging but considered by the genetics course instructor to be within the capability of these second-year biology students. In fact, she incorporated these articles into her lectures after the final paper deadline.

A counter-balance protocol was used in which one half of the students analyzed paper A first while the other half analyzed paper B first. This was done to minimize any inherent effects of the differences between the two articles being analyzed on the outcomes of the study. For the second phase of the assignment, they switched so that each student analyzed the paper they had not previously read in the first phase. To determine which paper each student should read first, the university's LMS was used to randomly split all students enrolled in the course into two groups. The randomization process of this LMS consists of two consecutive *PHP* functions: a "seed random number generator" or srand function followed by a shuffle protocol that ensures truly randomized samples.

Each group was assigned to read and critique only one of the two primary research articles in preparing their first analysis paper. Students had one week after the first paper assignment was posted in the course's LMS site to read their assigned article, analyze it and submit their first written analysis paper (see Figure 1). Two weeks after

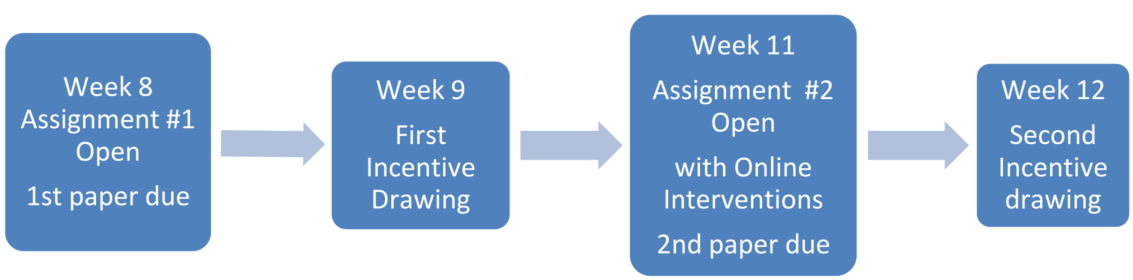


Figure 1: Timeline for study interventions, assignments and incentive drawings

the first paper deadline the students were assigned to read the other scientific study article chosen by their professor (the one they had not previously read) and prepare a second scientific study analysis paper on this second experimental study. At this time, students who consented to participate in the CA study also received instructions on how to access one of three online CA skill interventions. The second analysis paper was again due a week after all of these instructions were issued. The study participants had access to their assigned online intervention for the entire week leading up to the second paper deadline.

Key CA elements. The purpose of the first analysis paper was to provide baseline data on the CA skill level of each student. The assignment instructions clearly delineated the need to address within their analysis papers five key CA elements as follows:

1. Identify and focus on the main purpose and point of view of the study.
2. Identify and consider the most important scientific information provided and the key scientific concepts that must be understood in order to analyze the study.
3. Identify and consider conclusions drawn by the author(s) and the main assumptions underlying them.
4. Identify the main implications if the findings of this study are rejected or accepted and take a clear personal stand on the validity of the study results.
5. Analyze the credibility of the sources used by the author(s) in researching the topic of their research study.

These key elements were also part of the CA rubric the researcher had previously created and validated for her pilot study. Aspects of these key CA elements were also incorporated into the online educational interventions designed for this study.

Primary research articles and assignment instructions, differing only in the specific paper link were embedded at the end of an online lesson feature of the course's LMS website. The assignment instructions can be reviewed in Appendix B.

Prior to the introduction of the Scientific Paper Analysis Assignment, all of the students enrolled in the course had also been randomly assigned to one of three different treatment groups using the pre-set command within the university's LMS, as previously discussed. The students who had consented to be part of the study were then granted access to one of the three variations of online learning objects depending on their previously-assigned treatment group.

For the second and final phase of the CA study, all students enrolled in the course were assigned to read and analyze the other paper (either A or B) that they had not analyzed during the first CA study phase. Those participating in the study were also sent instructions on how to access one of the three online educational interventions: a simple game, a video or a tutorial. These three learning objects, with varying levels of interactivity, provided students with different perspectives on the five key CA elements to include in their written study analysis papers. Students had the entire week to read their second assigned paper, access their online intervention and write their second scientific study analysis paper.

The researcher sent an e-mail to all of the CA study participants with the link for accessing their assigned intervention and a brief worksheet for logging the time that they spent reviewing and interacting with their intervention during that week. A copy of the logging worksheet was also embedded in the introductory LMS site for each intervention. The students were asked to reflect on their interactions and complete their log by providing comments about their experiences and impressions while interacting with their specific online intervention. They were also asked to complete the CA study post-test in order to have their names included in the second incentive drawing.

Subject selection and description. The subjects of this study were second-year students enrolled in a 100-level genetics course that served as their second college biology course during the fall semester of 2014. It was expected that almost all of them would be in their late teens or early twenties (i.e. of traditional college age) and would be planning to major in biology or in an engineering field that was closely related to biology such as bioengineering or biopharmaceuticals.

The class met three times per week for lectures led by a professor and once a week for a recitation period which was led by a graduate student teaching assistant (TA). Their advisors in the biology department also strongly urged each student to concurrently enroll in one of two related laboratory courses that met once a week.

The professor leading the genetics course holds a PhD in biophysics and had prior teaching experience both as an undergraduate and graduate TA as well as a visiting lecturer at a large university in Southeastern USA. She was new to the institution, but had recently completed a post-doctoral fellowship in which she received extensive pedagogy training designed to help newly-minted science professors succeed as educators while also enhancing science education itself by having them design and use highly engaging teaching methods. She readily agreed to include what she called the "Critical Thinking Papers" in the syllabus for this genetics course, as long as the researcher agreed to grade all of the student papers. To increase the likelihood that the students would take this assignment seriously, the lectuere allotted 50 points for each of the two papers the students would complete. This made the overall assignment worth a total of 100 points or 10% of their overall course grade.

The researcher created a lesson within the university's Moodle-based LMS entitled "Analyzing a Scientific Paper." This lesson included a pre-test to collect information about the students' academic preparation for the course and their current level of knowledge about primary scientific research papers (see Appendix C). A link to an electronic consent form to become part of the CA study was also embedded in the written introduction to this lesson, but it was clearly stated that completion of this form and participation in the study itself was optional and would not affect their course grades. However, the introductory material also clearly explained that only study participants would be granted access to one of the three new online learning objects being designed specifically to help enhance CA skills. Using the lesson feature of the LMS was meant to ensure pre-test completion because a verification code needed to access links to the assigned scientific papers and the instructions for the scientific paper analysis assignment was available only at the end of the pre-test. These instructions also included descriptions of the five key CA elements that should be included in written analysis papers and were incorporated into the CA rubric (see Appendix A) used in this study.

Recruitment of study sample. During the third week of the fall 2014 semester, the researcher sent an e-mail to all students enrolled in the genetics course to introduce the study, explain the online interventions, and encourage students to participate on a strictly voluntary basis. Students were encouraged to volunteer for the study by signing a consent form (see Appendix D) that would allow the researcher to copy and use their analysis papers in the rating process. The e-mail also explained that as an incentive to participate, the researcher would hold be holding in-class drawings for one of six $25 gift certificates to the university bookstore. The first drawing for three of the gift certificates was held at the beginning of the first class after the deadline for submitting the first scientific analysis papers. The researcher's e-mail explained that only those students who chose to be in the study and also submitted this first paper on time would have their names included in the pool for these gift card drawings.

This introductory e-mail also stressed that neither the course lecturer nor the TAs would know who was or was not participating in the CA study so that this personal decision would not be able to impact their course grade. The researcher explained how confidentiality and anonymity of participants would be maintained throughout the study, even during the blind paper rating process as described below.

Confidentiality and anonymity. The researcher created a confidential participant database to use in tracking and analyzing each individual's study data that was available only to the researcher. Each student analysis paper was labelled with only student code number assigned by the researcher. No identifying information beyond the assigned code was visible on any hard copies of the papers themselves. Information on whether the paper was the first or second attempt by that particular student was also removed to ensure a totally blind rating process. The researcher assigned a separate course grade for completion of the analysis papers prior to encoding the participant papers and prior to holding any of the CA skill rating sessions.

To ensure that study participation or non-participation did not impact the students' course grades, none of the results of this study were shared with anyone from the biology department until the following semester, well after the fall semester grading deadline. Even then, only aggregate group data would be released from this study and included in any subsequent reports or publications, making it impossible to identify any single student's participation or CA skill scores.

Educational interventions

Three different online learning objects were originally created specifically for this study to help build understanding of the five key CA elements. The first intervention did not work out, helping to create an accidental control group who did not receive any information about the assignment other than the written instructions. The CA results of this group were also compared to those from the remaining two online interventions.

Expert CA movie.

The researcher asked three full professors at the university to review a copy of the CA rubric and gained their permission to videotape them individually as they discussed how they would use the key CA elements found in the rubric to guide their analysis of primary scientific articles. Each professor (two men and one woman) shared their thoughts on the key CA elements they use in analyzing scientific papers but none of them directly addressed the CA elements during their video-taped sessions. They did, however, provide rare insight on how seasoned scientists judge the quality of scientific research studies. The researcher used common themes from twenty-five minutes of recorded interviews to create a five-minute video montage of their remarks entitled "Determining the Credibility of Scientific Research Articles."

Figure 2: Scene from video entitled "Determining the Credibility of a Scientific Research Article"

This movie contained information related to the CA elements without explicit directions on how to analyze primary scientific research. Therefore the genetics students were challenged to transfer the information provided by these scientists to the task of preparing their own written analyses of the assigned scientific research study article.

Basic Game.

The game was the most challenging intervention to develop. The researcher quickly discovered that creating games requires more programming skill than she had anticipated. She did not have the expertise needed to do the concept justice for this particular study. Another student within the college of education who teaches computer programming at a regional high school was willing to allow his students to take on this project. Under his supervision, five high school students used the researcher's concepts and storyboards, along with open source images to create a simple, interactive game entitled "Scientific Proof." In this game, aspects of the CA elements are presented as requirements for successful scientific research grant proposals.

The researcher wanted to incorporate authentic opportunities to apply the key CA elements into the game. Therefore, when creating the concept for the game, she decided that the quest for research money would make a more compelling story (Broussard, 2012; Malliarakis, Satratzemi & Xinogalos, 2014) than would a quest for CA elements to include in a written analysis paper. The high school teacher who oversaw the development of the game liked this concept as well, as did the high school students who made it come to life and were only a few years younger than the target audience.

The game opens with a young scientist holding up a flask of a newly created elixir and wondering how he can obtain more money to study it further (see Figure 3). A female Lego© scientist serves as a mentor in this game, asking the young scientist (and therefore the students) a series of multiple-choice questions designed to help him prepare a proposal for a scientific research grant. The series of multiple-choice questions that she provides were designed to help the scientist (and students) think about the important elements needed for successful funding proposals without being too technical.

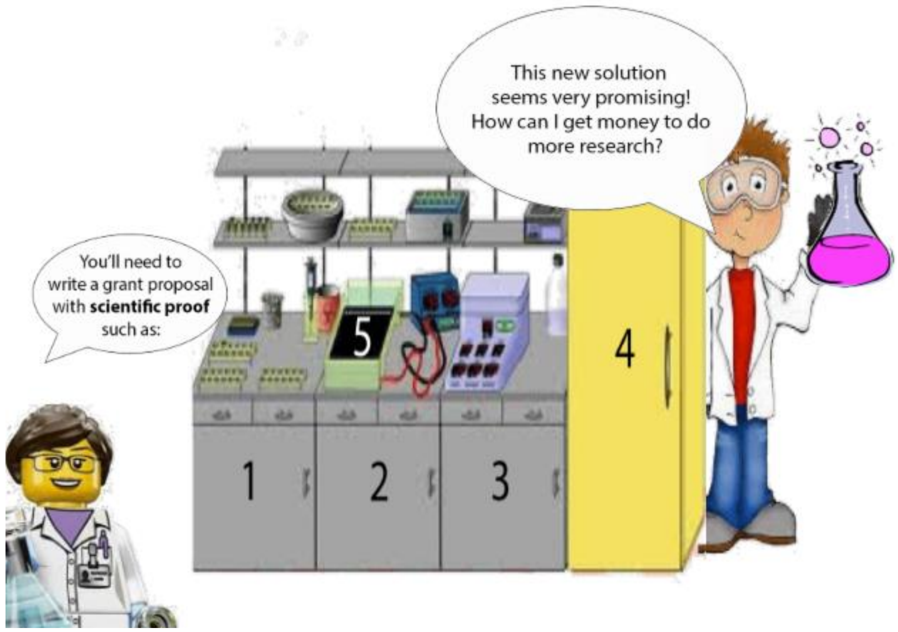


Figure 3: Opening scene from "Scientific Proof" game

Students playing this basic game had to open five cupboards, each corresponding to one of the key CA elements from the rubric that would be used to score their scientific analysis papers. This would display a multiple choice question with three.

to four options. When a student chose an option that was not correct, the option would simply disappear. If they chose the correct option, they would advance to a slide with positive reinforcement about the importance of that particular aspect in obtaining scientific grants. Each correct answer also earned the young scientist a stack of cash so that by the end of the game, the young man has enough money to fund the additional research needed to further improve his new discovery or scale up production.

Playing this game again challenged the genetic students to transfer the concepts presented to their assigned task of preparing a written scientific study analysis paper. In this way it was fairly similar to the expert video in the level of cognitive challenge and therefore differed from the video mainly through its playful images and by the higher level of interactivity that it required.

Each learning object had a worksheet embedded in their introductory LMS page along with a link to an online CA post-test (see Appendix E) as shown in Figure 4. The students were asked to note the time they spent accessing their assigned learning object, track it on the provided worksheet and then submit the completed worksheet to the researcher up to a week after the final paper deadline. The logs submitted by the students were later compared to the actual time spent by the students as tracked within the LMS.

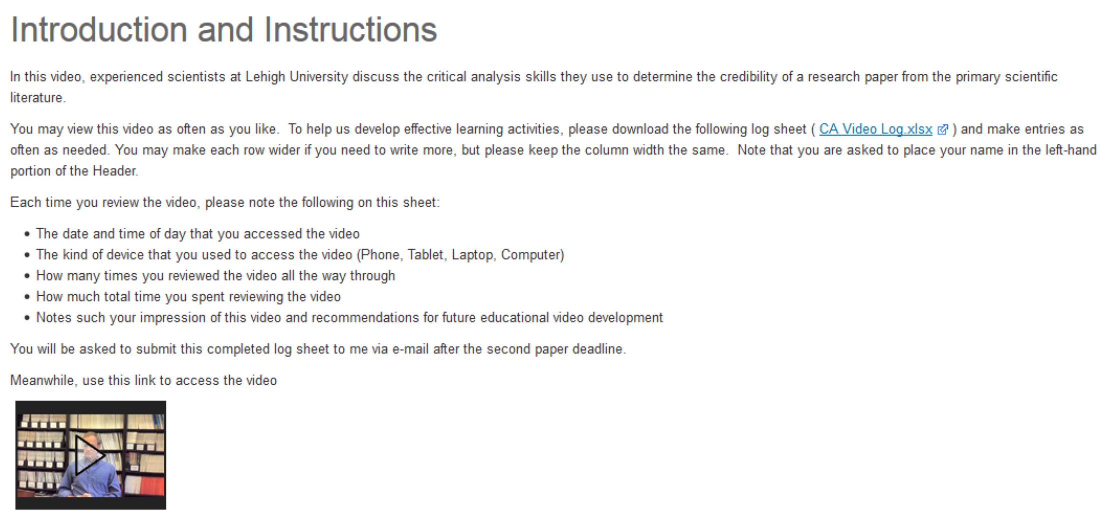


Figure 4: Sample introductory page for the Expert Movie within the Canvas© LMS

The CA post-test contained items that asked the students to share their opinions of the online interventions, how helpful they were for completing the scientific study analysis assignment and how they could be improved for future instructional purposes. The students' feedback is included in the fourth chapter of this dissertation.

Collecting and rating study artifacts

The scientific study analysis papers were required assignments for all students in the course, whether they decided to participate in the CA skill study or not. All students submitted both analysis papers within the LMS site for the genetics course. While the course instructor had the ability to access the papers within the submission inbox of the Scientific Study Analysis Paper assignment in the course's LMS website, she was unlikely to do so, since the researcher was responsible for grading all of the papers. Even if the course lecturer had accessed the inbox, she would not have been able to distinguish who participated in the study because all of the students were listed there by name only with no reference to whether or not they had signed the study consent form. The researcher assigned random code numbers to each individual paper while creating a key that matched the papers to the students who wrote them. This key has been kept secure at her home and will not be placed anywhere that it could be accessed by others.

Anonymizing the student papers

The researcher opened one student paper at a time in the assignment's electronic submission inbox. She removed all identifying information and any references to the current paper being either the first or second one written. She replaced the student's name with their assigned code number. She then printed a single copy of the now anonymized paper before closing it without saving any of the changes she had made to it. This single copy was then photocopied so that each rater could have an individual copy of each paper that they could mark up as needed during the closed-door rating sessions. None of the course instructional staff (lecturer nor TAs) were involved in these processes.

Description of the CA assignment rubric

The raters measured and quantified the CA skill level of the students by using an assignment-specific rubric that was previously designed and tested by the researcher. This rubric was created by adapting and then merging two different resources. A generic scale and CT descriptors developed by the Washington State University Critical Thinking Project were combined with the key CA elements derived from the Foundation for Critical Thinking (Paul & Elder, 2003) as previously discussed. These elements were used as the criteria of the rubric which were shared with the students as part of the instructions for the Scientific Study Analysis Assignment.

The original CT scale developed by WSU defined levels for judging student demonstration of CT skills within generic assignments. The researcher adapted these to create a distinct six-point scale as follows. The scale starts with a score of "0: when mention of a particular CA element is totally absent from the student's analysis paper. Assigning a level of one for a particular element indicates "minimal" inclusion or development of that element in the student analysis paper. For the rest of the six-point scale, two indicates "emerging," three indicates "developing," four indicates "competent," five indicates "effective" and a score of six indicates "mastery" of that particular criteria (CA element). Since the rubric contains five identified criteria (CA elements), an overall CA score of twenty (average score of four times five criteria) would be considered a minimal level of competency. A score of six on each of the five criteria would yield the rubric's maximum overall CA score of thirty, denoting CA skill mastery of the highest level.

The WSU descriptors, written with input from hundreds of teachers over several years as part of a program funded by a large grant from the National Science Foundation were intentionally vague so that they could be easily used for almost any subject. In splitting the scale elements into six, independent parts, the researcher had to split the three descriptors into six that would allow for the kind of discrimination needed for scoring the papers and statistical analysis of quantitative differences rather than only qualitative ones. For the previous study, each of the five criteria (key CA elements) was placed on a separate page and merged with the new, more discriminating, six-point scale to create a CA rubric specifically tailored to the scientific study analysis assignment (see Appendix B). The researcher successfully used this new CA rubric to demonstrate differences in CA skills between treatment groups in a pilot study using a similar scientific study analysis assignment (Adams & Columba, 2014).

Rater selection

Two professional women from the community college where the researcher had completed the pilot study agreed to serve as raters for this study. Both had a strong interest in CT and were themselves excellent writers. One is a college English professor with a Ph.D. in Literature. She has attended conferences of the Foundation for Critical Thinking and shares responsibility for providing faculty development CT seminars at a mid-sized community college. The other recently received a Ph.D. in education and is a seasoned administrator involved in community education program development. She serves on the college's curriculum committee and has a special interest in ensuring a well-rounded curriculum that helps students develop critical thinking, leadership and other valuable lifetime skills for the workplace.

Rater Training

After the second paper deadline, the researcher gave both raters two anonymized papers from the pilot study that she had previously run at the community college. She also provided them with copies of the latest version of the CA rubric (see Appendix A) review. The raters were asked to use this rubric to rate the CA elements within the paper and return their ratings to the researcher before their first training meeting.

At the beginning of the first training session, each rater again assigned CA scores to copies of the pilot papers that they had previously rated. The researcher statistically compared each rater's earlier CA scores with those of this second round to determine initial internal validity.

The researcher then reviewed the purpose of the CA study and the evolution of the CA rubric. She also discussed the assignment that was given to the students and provided a brief summary of the online interventions. Then all present discussed the CA rubric and any questions they had from scoring their first set of qualifying papers.

The raters were then given a new paper from the pilot study and asked to score it together as a group, discussing specifically how they had decided which level of proficiency to assign for each of the five CA elements demonstrated within this paper. The two raters then individually rated more sample papers and their rating scores were compared and analyzed to determine the level of inter-rater reliability (IRR). As long as the rating scores from the two raters correlated at least 70 percent or higher, they could continue to rate experimental papers (Jonsson & Svingby, 2007; Stemler, 2004). If IRR was ever found to be lower than 0.70, then all of the raters were re-trained with an emphasis on building consensus for each of the five key CA elements. They then had to rate additional qualifying papers from the pilot study until acceptable levels of reliability were again restored. The rater testing and training process described here (Saxton et al., 2012) would then be repeated to ensure maintenance of acceptable levels of both intra- and inter-rater reliability throughout the rating process, further validating the use of this rubric for use in CA skill studies (Saxton et al, 2012).

Paper rating process

The researcher gave each rater an anonymized and coded copy of the scientific study analysis papers submitted by all of the students participating in the CA study. Unique codes were randomly assigned to each individual paper. This meant that even though each student had written a pair of papers, each individual paper would proceed through the rating process independently, without being tied to its match. This also ensured that an individual's first and second paper received objective ratings without regard to the order in which they were written. The raters reviewed each individual paper and used the CA rubric to assign scores from zero (absent) to six (mastering) for each of the five CA elements within each individual analysis paper.

Data collection.

These individual scores were then summed to determine each rater's overall CA score for that paper. The overall CA score along with the individual elemental scores from each rater were then averaged and potential differences between the treatment groups were determined via a mixed ANOVA with repeated measures using SPSS.

Three online interventions served as the main independent variables (IVs). The effects and interactions of all of these variables were determined using mixed designed ANOVA. This particular statistical method was chosen because it supplied not only the main effects between the groups, but also any interaction effects that may have been present between the independent variables of the study.

This provided four potential sources of quantitative data for each participating student: (1) baseline CA skill level scores on their first scientific analysis paper; (2) post-intervention CA skill level score from their second paper (3) self-reported CA skill assessment from the CA pre-test, and (4) self-reported CA skill levels from the CA post-test (see Appendices B and C). Qualitative data was collected from three sources: (a) pre-test information on demographics and completed coursework (b) postscripts on their personal writing process submitted in their intervention logs, and (c) post-test evaluations of the online interventions and the CA study (see Appendix E). Results from the CA post-test were compared to those of the pre-test for all students who completed both.

Data from all of these sources were compiled and analyzed for each student to ensure triangulation of data. The results this analysis and of the interactions between the independent variables is discussed in the next section of this dissertation.

Statistical power analysis.

The researcher had previously completed a study on the effects of three different live educational interventions on CA skill development by second-year college students completing the same type of assignment at another institution. The significant differences in CA skills found between the intervention groups in that study had been quite large with F-values of 5.54 and 12.67 for the respective differences in student coverage of the implications and of the quality of the resources used in planning the experimental studies that the students analyzed ($N = 50$, $\alpha = 0.05$). However, there had been a lower than expected yield of data in that study. Only 35% of the potential student subjects who were enrolled in participating biology course sections had both consented to being part of the study and then also submitted a suitable paper that could be used for measuring their demonstrated CA skills. Two factors identified to explain the high attrition rate of interested subjects in that study had been the high course drop-out rate and the low point value apportioned to the scientific paper analysis in those courses so that if the students were very busy, other academic tasks with more impact on their course grade took precedence over the paper. Neither of these were a factor this time, yet the researcher still anticipated about half of the students (75 total) would decide to participate in this study.

Chapter 4

Results

This study took place within a group of 142 students taking genetics as their second biology core course at a small, private college in the Northeastern United States. The students enrolled in this genetics course were primarily second-year students who successfully completed introductory biology during their first year of college and were enrolled in introductory genetics as their second biology core course. However, only 31 of the 135 students who took the CA Study Pre-test had completed 39 credits or less as expected for sophomore standing. In fact, 12 students who took the pre-test (9%) had already completed over 100 college credits which corresponds with senior standing.

The professor/lecturer for this course was trained as a scientist and had just completed a unique post-doctoral program that included extensive instruction in science pedagogy. She was enthusiastic about trying new educational approaches and met with these students twice a week for 50-minute large-group lectures with occasional in-class group learning activities. The students also met once a week in smaller recitation groups led by graduate student teaching assistants (TAs). In addition, most (75%) were also enrolled in one of two highly recommended, but not mandatory, one-credit laboratory courses designed to accompany the genetics course. These laboratory courses challenged students to apply the lecture material.

Recruitment of study subjects

The researcher first sent an e-mail explaining the study and then visited the class during the fifth week of the fall semester, two weeks before the online assignment opened for the students within the university's Learning Management (LMS). She explained that the purpose of the critical analysis study which was to help stimulate the development of critical analysis (CA) skills in science instruction at the university. She then introduced the study incentives and encouraged them to participate in the CA study.

Of the 142 students enrolled in the course, a total of 69 (or 49 %) completed the online electronic consent forms, thereby agreeing to participate in the study. Forty-two of those participating in the study (61%) were women and the remaining twenty-seven (39%) were men. This gender representation is more reflective of the university's College of Arts & Sciences which was offering this particular course than of the university as a whole. For example, women account for 65 percent of the undergraduate students in the College of Arts and Sciences but only 44 percent of all undergraduate students enrolled at the university.

All of the students enrolled in the genetics course were randomly assigned, again using the PHP functions of the LMS, to one of two "paper" groups. Group A was to read and analyze Paper A first while Group B was to critique Paper B first.

As mentioned previously, the researcher used the lesson feature within the LMS to ensure that the pre-test data was representative of as much of the class as possible. A verification code which the students needed to access the primary research papers and the specific instructions for analyzing them was found only at the end of the lesson within the university's LMS entitled "Analyzing a Scientific Paper." However, a dozen students (9% of the class) completed the pre-test without noticing the code at the end until they had closed out of the pre-test. The researcher had set the LMS to allow only one attempt at the pre-test to prevent students who had already taken it from confounding the data with further attempts. However, this setting also prevented them from getting back into the pre-test only to obtain the verification code.

When these few students alerted the researcher to this problem, the researcher first checked the system to ensure that the student had truly completed the pre-test before sending out the verification code. However, since both she and the course instructor had received a disruptive amount of e-mails on this issue, the researcher decided to modify the ending of the pre-test so the importance of the verification code for continuing on with the next steps of the lesson was more obvious to the students. She also decided to change the settings to allow up to three interactions per student.

The electronic study consent forms as well as the pre- and post-tests for this study were administered through a web-based survey software program called *Qualtrics* that was licensed by the university's technology department. The *Qualtrics* dashboard showed that 143 students had opened the "CA Pre-test and Academic Prep Survey" but only 135 (or 94%) of these students had actually completed it. Upon review of the log of responders, it was found that eight students had completed the pretest twice and one had taken it three times, seemingly in an apparent effort to locate the verification code. Only the information from their original attempts were used in the study data analysis.

Five of these pre-test repeaters also signed multiple consent forms to participate in the CA study. Three of these students sent e-mail messages to the researcher stating that that they had "signed the consent form to be in the study" specifically to gain access to the scientific paper analysis study instructions. In fact the *Qualtrics* dashboard indicated that there were 94 signed consent forms, however upon further investigation of the detailed log, it was discovered that there were only 69 individuals who had actually consented to participate in the study, some multiple times.

Note that most of the students (91 %) enrolled in the course followed the initial instructions and had no trouble completing the pre-test and accessing their designated paper and the assignment instructions. However, the researcher and the course instructor were kept busy for a couple of days helping the other nine percent of the students who ran into various difficulties with the process. During the resulting deluge of e-mails from this individuals about assignment access problems, the course professor/lecturer decided to provide the verification code to some of the students who had not yet accessed the pre-test. The survey analytics showed that approximately the same number of students had unsuccessfully tried to access the next step in the lesson using variations of the pre-set verification code which was an encapsulation of the central dogma theory of biology meant to help reinforce what the students were learning in the genetics course.

Study pre-test results.

The online introductory lesson had been set up to steer all of the students to the study pre-test and so gain baseline data. This study pre-test (see Appendix C) collected demographic information about the students in the course including their gender, age, race, college major, number of college credits completed and how many relevant college courses they were currently taking or had taken in the past. Science, engineering, writing and history courses were all considered relevant to this study since they have all been linked to enhanced CT skills as discussed in the literature review outlined in chapter two of this dissertation.

A total of 135 students completed the CA Study Pre-test and Academic Preparation survey available through the *Qualtrics* web application as previously discussed. Over sixty percent of these students were majoring in one of the expected fields with (37%) majoring in biology and another 27% majoring in bioengineering. The rest of the students were either majoring in chemistry (12%), molecular biology (4%), medicine (4%) or IDEAS (4%), a four-year interdisciplinary honors program at the university. However, 12% of the respondents had chosen not to provide an answer to this pre-test question.

The pre-test confirmed that almost all of the students taking this course were traditionally-aged college students. All but three of the 135 respondents (98%) were no more than 22 years old with an average age of 19.5 years old for the entire class.

In the academic preparation portion of the pretest, it was revealed that the average grade point average (GPA) of the men who decided to participate was 3.32, only slightly higher than the 3.27 GPA earned by the women in the study. Forty-six men and ninety-two women reported completing at least two science courses and fifteen students (six men and nine women) reported completing at least one engineering course. However, it appears that at least some of the responders did not notice the shift from "credits earned" in the first of these questions to "courses taken" in the second. Unexpectedly high numbers of courses were reported including one student's response that they had completed 79 science courses. The open-ended nature of the question paired with the fact that it directly followed a question about total college credits completed confounded the answers so that these variables could not be trusted for further statistical calculations.

The other academic subjects that were predicted to impact the results of this study were history and writing. Thirty eight survey responders (16 men and 22 women) had taken at least one college history course, although again it was uncertain if a response of "3" meant three courses or one three-credit course. Only fifteen students (11%) reported taking at least one college writing course. This lower than expected amount may have been because 67 students (47%) reported that they had received college credit for the Advanced Placement English courses they had taken in high school.

The rest of the questions on the pre-test revealed that the vast majority of the students knew exactly what a review article was and how it differed from a primary research article, even though five people responded that they did not know the answer to this question. Notable responses to this question included the following:

> A research article is written by the scientists, who performed the research, while review articles are from a secondary source.
>
> A review article is a survey of pre-existing primary research articles. A primary research article is an analysis of new experimental data.
>
> Review articles aren't a primary source and are often written about primary sources – research articles.

The students were more uncertain of where the study data would be found within a scientific paper with the most popular answer being in the results section as can be seen in the word cloud generated by *Qualtrics* and pictured in Figure 5. The size of a particular word or phrase within a word cloud is directly proportional to the frequency ot of that particular answer within a survey or a paper.

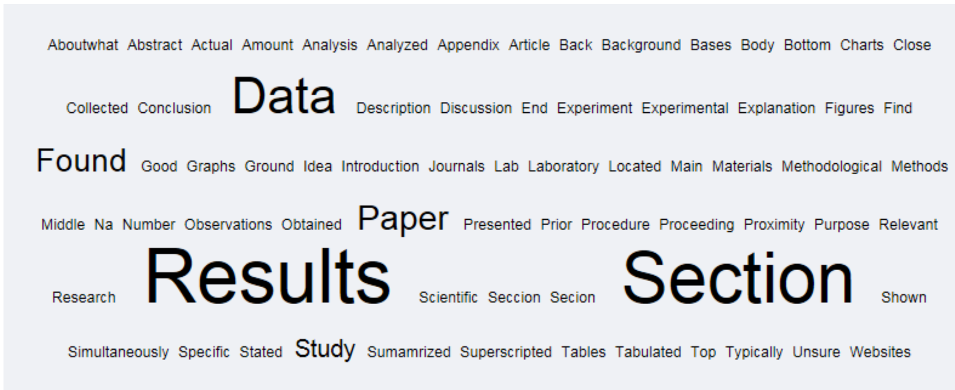


Figure 5: Word cloud compilation of student answers to the question of where data would be located within a scientific paper.

Almost all of the students knew that the sources used by the researchers in planning their study could be found in the reference section as seen below in Figure 6.

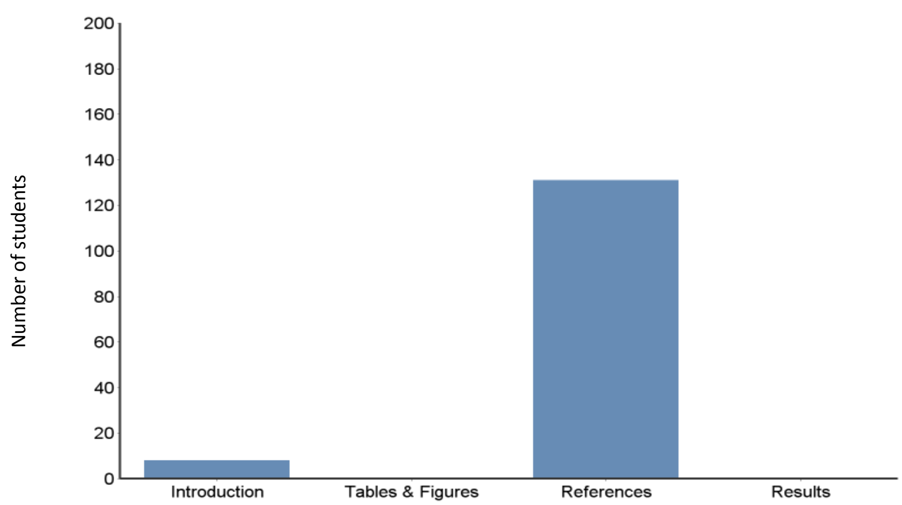


Figure 6: Compilation of student answers to the question of where the sources used by the researchers/authors in planning a scientific study would be found within a scientific research paper.

The last section of the pre-test asked students to rate their own level of competency in CA skills including scientific literacy, information literacy, research, critical science reading and science writing using a Lickert scale. The scale for these self-reported ratings could range from one for very poor to five for very good. Most of the students (mean of 35%) rated themselves fair or good (mean of 64%) in all five areas as can be seen in Figure 7.

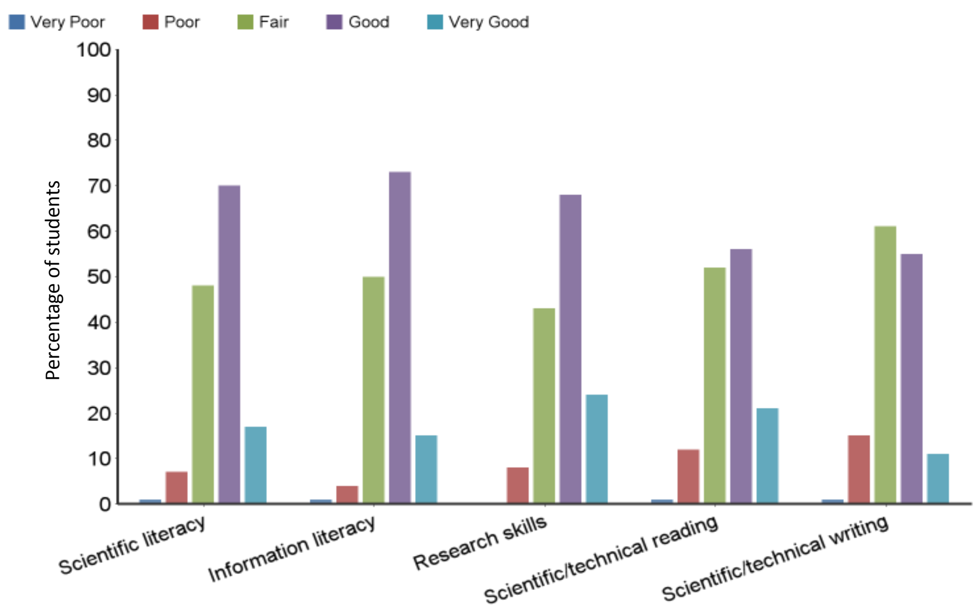


Figure 7: Student ranking of their own CA skills related to analyzing scientific papers (from the study pre-test)

Description of the Assignment

The last part of the lesson within the LMS entitled "Analyzing a Scientific Paper" provided the specific assignment instructions and library links to one of the two assigned papers. All of the students in the course were randomly split into two groups by the LMS. Each group was then assigned to read one of the two scientific papers. The students in Group A were to analyze Paper A first while the rest of the students who were all in Group B were assigned to read and analyze Paper B first. Written assignment instructions that included the identified CA elements were provided to all of the students enrolled in the course (see Appendix B).

Primary research articles

The course lecturer chose two primary research articles from the field of genetics for the students to analyze in the scientific study analysis assignment. She chose these specific papers because they complemented material she planned to present in lectures near the end of the semester.

Paper A summarized a comparison study of DNA methylation patterns in early-stage embryos from five different species including humans (Fulka et al., 2004). Methylation of various parts of the embryonic DNA is nature's way of turning on and off access to the genes needed for producing the proteins responsible for normal embryonic development. The experimental procedures outlined in this paper ultimately kill the embryos which is why the work was done only up to the blastocyst stage in the human embryos, which ends approximately one week after fertilization. Live research studies of human embryos beyond this point are highly controversial for ethical reasons and currently restricted in the USA. This level of detail would have been expected of students of the mastering level for the scientific information criterion of the CA rubric. Many in the scientific community agree with these researchers and their procedures have been replicated in more recent studies as evidenced by the 122 times this paper was cited in the *Web of Science Research Database* over the ten years since it was first published.

Paper B was a study of the effects of a particular gene (AUTS2) implicated in the development of autism spectrum disorders. This paper outlined a unique protocol for determining the specific traits encoded by any gene. This protocol includes knocking out the gene to determine the overall effects of not having it activated on the characteristics of the test organisms as well as systematically adding back specific aspects of the isolated gene to see if the changed characteristics noted can be restored in other organisms.

First scientific study analysis paper

After the students read their first assigned paper, they were to prepare a written analysis and submit it through Turnitin© a plagiarism prevention software program that compares their papers to an extensive online database. The researcher set the paper link to allow multiple submissions so students could directly monitor the originality of their drafts and learn through repetition how to properly cite. However, Turnitin© did not always return results in a timely fashion and appeared to significantly slow down when there were more students submitting, as happened just before the paper deadline. Several students wrote e-mails to the researcher about this issue just before the-first paper was due, venting frustration that they received messages through the assignment link that their "originality score" would not be available to them until after the first paper deadline.

Another source of student angst was the stipulation that the papers be 700-800 words in length, excluding the title and references. This requirement proved to be difficult for many students. Only a few of the submitted papers were shorter than 700 words. However, many students struggled to limit themselves to only 800 words. This issue generated numerous e-mail messages to the researcher who reminded them that in the real world scientists deal with similar size limits when writing grant proposals for research funds as well as primary research papers for submission to academic journals.

The researcher had agreed to grade the papers from all of the students enrolled in the course whether or not they participated in the CA study. However, she and the course instructor had also agreed that providing grades or any other kind of feedback to the students about their first analysis paper would make it virtually impossible to differentiate the effects of the online interventions from the effects of the feedback. The researcher explained this in person to the students when she visited the class to hold the first incentive drawing for three gift certificates to the bookstore. The course instructor reinforced as needed that this lack of feedback was a necessary aspect of this educational study. However, the students were not happy and many expressed their concern that this would negatively impact their course grades since together these papers were worth 100 points or 10% of their overall course grade.

Originally, there was to be only one week between the first paper deadline and the opening of the second paper assignment and interventions within the LMS. However, the course instructor and the researcher decided that major social events on the weekends following the first paper deadline (Parent's weekend and Halloween) would make it difficult for the students to find the time to adequately access the study incentives and prepare their written assignment papers. Therefore, the opening of the second scientific analysis paper assignment within the course's LMS website was postponed to the eleventh week of the semester.

For this second assignment, students were to read and analyze the other primary research paper that they had not previously read. Again they were given a week to prepare their second written analysis paper. Many students submitted their second papers earlier, some within a day or two of this assignment's first opening. The researcher thought that this was done to more fully utilize the self-monitoring capabilities of Turnitin©. However, in hindsight it was probably more due to the fact that midterms were also held during this week which will be discussed more in chapter five.

During the week all students were to read their second papers, those who had consented to participate in the CA study were also granted access to one of three different online learning modules related to analyzing the credibility of primary research papers as described below. The students submitted their written analysis of their second primary research article by the end of the eleventh week of the fall semester. The researcher was not able to complete grading all of the papers until after the course final which was stressful for all involved.

Online Interventions

Prior to introducing the study to the students, the researcher used the PHP functions within the university's LMS to randomly divide all of the 142 students into three treatment groups. As mentioned before, a total of 69 students chose to participate in the study and by chance, exactly 23 students from each of the previously identified groups signed study consent forms creating equal sample sizes for the three interventions. To prevent cross-contamination, the researcher set up separate mini-courses in a free LMS program called Canvas©. Only the students assigned to each intervention were enrolled in the corresponding mini-course. The opening page for each course included a welcome message, an embedded spreadsheet (log) for students to use to record their interactions with the learning modules and a link to the intervention itself.

On the same day that the second scientific analysis paper assignment folder was available to the students within the course's LMS site, the researcher enrolled the students in the Canvas©-based course for their assigned intervention group to generate a welcome e-mail from this LMS to each of the students. She also sent an e-mail to the study participants advising them of their intervention group, providing a direct link to their course and alerting them to the fact that they should be watching for a welcome message from Canvas© with more detailed information on accessing the Canvas© course that contained their assigned intervention.

The online interventions in this study provided information related to analyzing the credibility of primary scientific research papers, without explicitly explaining how to prepare a written scientific study analysis paper. Each online learning intervention was designed to be completed in approximately five minutes. Recent evidence indicates that the adult attention span may be much shorter than originally expected (Sousa, 2011). This has led to an evolving "best practice" in online learning design of providing short educational vignettes that better match this limited attention span in an attempt to enhance learner engagement and lead to better learning (Khan, 2012; Sousa, 2011).

The week before the second analysis paper deadline, the students in this study were able to freely access their assigned intervention. They were also asked to read the primary research paper that they had not read before and prepare a written analysis of it by the end of the week. Both analysis papers were then assessed for CA skills by trained raters using the CA rubric as previously described.

A total of 21 students or 91% of Group A accessed the basic game intervention prior to submitting their second paper. One student revisited the game site two weeks later. Their interaction times with this online game ranged from just over two minutes to over five hours with an average of 31.81 minutes spent on the game (see Table 1).

While 14 students from Group B accessed the CA video website, one student who accessed it did not open the video at all. The other 13 students (57% of Group B) opened, and presumably viewed, the video. The interaction times with the video ranged from two minutes to over two hours with an average of 38.28 minutes which was even longer than the average time spent with the game. All of the students interacting with the video did so only prior to the second paper deadline; none re-visited the site later.

However, as previously mentioned, a third online intervention did not work out as planned, leaving many students who had not accessed any online interventions. Since the researcher had the consent of these students to copy and rate their papers, she decided to use them as an accidental control group. This decision also allowed for a total student sample size that was large enough to have the statistical power to detect large differences in CA skill development among the experimental treatment groups. As outlined in the table below, there were still three treatment groups in this study: the basic game, the expert video and the control groups. In addition, the game and the video differed enough in interactivity to still be useful for attempting to answer the second and third research questions as previously outline in this dissertation.

Table 1

Study treatment groups

Group	Average Access Time	Gender	n
Basic Game	38.81 minutes	11 women 5 men	16
Video	38.28 minutes	8 women 5 men	13
Control	N/A	9 women 13 men	22

It should be noted that "originality scores" assigned by Turnitin© are really indicators of how much a student's paper resembles other student papers. In other words, it is really a measure of how much of the paper may have been copied. This means that lower originality scores are actually better because they signals less plagiarism and more original thinking and paraphrasing on the part of the students. However, similar phrases needed to explain technical information and using similar references will also be counted as possible problems, so teachers need to thoroughly investigate unusually high originality scores.

In this study, all of the originality scores were much higher for the second analysis papers submitted by the students. Some of the difference can be attributed to the fact that the first set of study papers had already been added to Turnitin's online database, increasing the likelihood of finding potential matches that could be flagged for possible plagiarism. However, about five students very high originality scores on the second papers that upon further investigation, turned out to be direct plagiarism from their classmates' papers. Only one of these questionably original papers was submitted by a study subject. Both this paper and its match from the first scientific paper analysis assignment were removed from further study.

Rating the student papers

After the two raters successfully completed the training program described in chapter three, they were eligible to rate the papers generated by the students in this study. However, the researcher continued to monitor internal and inter-rater reliability throughout the rating process by requiring the raters to score pilot study papers before each rating session. The raters took copies of the same three pilot study papers home with them at the end of each rating session. No more than twenty-four hours before the next rating session, they would rate these "qualifying papers" and electronically submit the scores to the researcher. The researcher analyzed the submitted data and if the raters' scores correlated at a level of at least 0.70 with each other and with their own past ratings of the same paper, they could start rating new experimental papers upon their arrived at the rating session. If not, the raters would discuss why they had assigned certain scores to the various criteria within the rubric for each of the three pilot papers. Then they would be given copies of three more pilot papers to rate. If the correlation on these new qualifying papers was acceptable, then and only then could they start rating experimental papers from this study.

The first rater training session was held on the last weekend in November. Two more weekly training sessions were held before the end of the fall semester. At the third session, the raters achieved both adequate internal and inter-rater reliability (IRR) scores

($\Gamma \geq 0.70$) while rating three different sets of qualifying papers. One of these sessions focused on the science within the assigned papers themselves since neither rater had a strong science background. There was a two-week break in the study rating process for grading finals and for the holidays. When the raters met again in late December, both internal and IRR on two additional sets of qualifier papers were found to be over the minimal acceptable score ($\Gamma \geq 0.70$). The raters were therefore able to end this session by rating their first set of four randomly-selected experimental analysis papers. The researcher randomly assigned these and all subsequent sets of papers until they had rated over fifty papers. She then began to ensure that at least half of the papers she gave the raters matched papers of individual students that they had already rated within the three final treatment groups (i.e. Basic Game, Video and Control groups).

Refinement of the CA rubric

Each rating session started with a validity check of the raters using at least three qualifying papers. When the raters did not match, the raters would discuss their reasoning for assigning the scores they had before then scoring another set of qualifying papers. Most of these re-training sessions were the result of ambiguous wording within the descriptors of the CA rubric itself, especially within the criteria dealing with the main purpose/point of view (criterion 1), implications of the study (criterion 4) and the references used by the researchers/authors of the assigned scientific articles in planning their studies (criterion 5).

When discrepancies occurred, the raters and researcher would discuss the specific wording although all of the criteria were at least slightly modified. that was leading to the different scoring interpretations. Once they understood why each rater had assigned a different score, they modified the descriptors to ensure a logical progression of CA skills within that particular criterion of the rubric. In this way, several rating sessions led to clarifying and refining the rubric itself. Each time this happened, only one round of papers was needed to restore acceptable levels of reliability.

Most of these descriptor clarification sessions involved ensuring a gradual level of competency for each criteria, which is denoted with a scale value of four. The raters and the researcher modified the wording so that a clearer, more logical progression was made for scores between three (developing), four (competent) and five (effective) within each rubric criterion. As the rubric evolved into a more discerning one, the reliability of the raters also improved so that IRR was consistently found to exceed the minimal level of correlation ($\Gamma \geq 0.70$) for group comparisons. In fact, IRR levels were usually found to be between 0.80-0.85 on the qualifying papers near the end of the paper rating process.

The researcher identified and set aside the experimental papers that had already been scored prior to these rubric modification sessions, in case they wanted to re-score those papers. Originally the researcher had wanted to maintain equal group sample sizes to ensure a more robust statistical analysis. Therefore, when all matching papers had been scored for a total of 39 students (thirteen from each of the three final treatment groups), the researcher calculated IRR using Cronbach's Alpha as suggested for this type of study (Field, 2013; Leech, Barrett & Morgan, 2008; Stemler, 2004). This first IRR analysis score for experimental papers ($\alpha = 0.854$, N=24) exceeded the minimal score needed for comparing study groups when testing across all 24 measurements and a slightly lower, still acceptable score ($\alpha = 0.779$, N=4) when calculated using only the overall CA scores for both papers.

The researcher then ran a post hoc power analysis to help determine how many study subjects would be needed to detect larger between groups effects such as the ones that were found in the pilot study. This second power analysis using a larger effect size ($f = 0.4$) and the same type 1 error rate ($\alpha = 0.05$) revealed that a total of only 51 subjects would be needed to reach an acceptable power level (0.824) for this study.

At this point in the study the researcher had a total of 138 papers from 69 participating subjects that were being blindly and randomly rated by the study raters. Since the new, stricter protocol for the rating process had significantly limited the pace of the paper rating process, the raters had scored a total of only ninety experimental papers by mid-March. However, when the researcher used her participant coding key to match the two papers written by individual students, she found that paper pairs had been completed for only thirty-nine subjects.

The ad hoc power analysis had shown that the statistical power of the study results would more than double if they could add another 12 students for a total sample size of 51 subjects. Given the need for additional power for the main study effects and the excellent initial IRR results of the completed paper matches as just described, the researcher chose to ask the volunteer raters to rate the additional papers needed to complete these matches. Rather than asking them to go back and review the papers they had previously rated in the early rating sessions, she felt that their limited time and patience with the protocols would be better spent gaining more statistical power for the treatment results. Ten of the last twelve papers the raters scored were written by students in the newly-defined control group that served as the third study intervention along with the video and the basic game.

Descriptive Statistics

Overall CA scores. The scale of the CA rubric (see Appendix A) correlates the level of inclusion of the CA element and sophistication of the accompanying discussion to that particular student's CA skills. The overall CA scores in this study ranged from a low of five ("minimal") to a maximum of 18.5 (almost "competent"). This means that none of the student papers in this study achieved an overall CA score that would be considered to be competent. This was not surprising given that the pre-test had shown that approximately one third of the students were only in their second-year of college and most of them had not yet taken any college writing courses.

The most unexpected finding was that the average overall CA scores for the first student papers were significantly higher than the average overall CA scores that the students received for their second papers ($F_{(2,48)}$ =16.92, $p < 0{,}01$, $\eta^2 = .261$). As Table 2 illustrates, the average mean overall CA score on the first paper was 11.41 (M =11.41, SD = 2.29, N = 51) and that of the second paper was 10.20 (M =10.20, SD = 2.13, N = 51). The IRR for Paper 2 (α = 0.760, N=10) was found to be adequate for comparison purposes. However, the IRR for the sum of the CA elements for Paper 1 (α = 0.652, N=10) was found to be only minimally adequate for comparison purposes (Field, 2013).

Table 2

Average overall CA scores

Measure	N	Mean	Standard Deviation	IRR (α)
Paper 1	51	11.41	2.29	.760
Paper 2	51	10.20	2.13	.652*

* Minimally adequate for comparison purposes

This is an average loss of overall CA skill demonstration of eleven percent between the first and second analysis papers. This means that overall CA skills were higher before the students received any type of experimental treatment. Dividing these average values by five to determine the average score per CA criterion, yields average criteria scores for both papers that are in the "emerging" range of the rubric (M=2.28 and 2.04 respectively for paper 1 and 2).

To assess whether the five CA elemental scales that were summed to create the overall CA score formed a reliable scale for the new sample size, Cronbach's Alpha was again computed. As previously mentioned, these scores needed to equal or exceed a score of 0.700 to be considered adequate reliable for comparisons of the main effects of the study (Field, 2013; Stemler, 2004). Again, acceptable values were found ($\alpha = 0.777$, N = 20) when all of the elemental scores were taken into account and also when overall CA scores from both raters were included in the calculations ($\alpha = 0.863$, N=24).

When IRR scores were calculated separately for the five CA elements, they ranged from a low of 0.337 for the reliability of the study conclusions and assumptions to a high of 0.830 for the credibility of the references as illustrated in Table 3.

Table 3

Inter-rater Reliability of key CA elements using Cronbach's Alpha

Purpose	Science Info	Conclusions/Assumptions	Implications	References
0.476*	0.607**	0.337*	0.726	0.830

* Inadequate for comparison purposes
** Moderately inadequate for comparison purposes

Further analysis of intra-rater reliability (internal validity) revealed low internal consistency for all values except for rater #2's total CA scores on Paper 2 ($\Gamma = 0.724$) and her scores when rating criterion 5 ($\Gamma = 1.000$). Despite the more stringent reliability protocol and excellent IRR correlations achieved throughout the qualifying process, the internal validity of the ratings of experimental papers was less than adequate for comparison purposes (Field, 2013; Stemler, 2004). This underscores the importance of continually testing not only IRR, but rater consistency as well. The researcher chose to continue to analyze the data for main effects, but any results must be considered tentative at best until such time as the papers can be re-scored by trained raters.

Main study effects

The researcher decided to analyze the three main effects of paper order, group and gender by using mixed methods ANOVA rather than MANOVA so that interaction effects between the independent variables could be more easily analyzed (Field, 2013). All of the assumptions needed for reliable ANOVA testing were met as described here, unless otherwise noted in the following notes of the results of each test. The assumption of the independence of observations was addressed by randomly dividing the students into groups, randomly coding the student analysis papers and then shuffling them so that they were rated in a totally independent, random order.

The assumption of equal covariances of the dependent variables and homogeneity of variance were met with observation of insignificant results of both Box's M test and Levene's test. Mauchly's test was used to test the assumption of sphericity of the data. A Shapiro-Wilk's test ($p > 0.05$) and visual inspection of their histograms, Q-Q plots and box plots showed that the paper scores were approximately normally distributed with skewness and kurtosis levels that also did not significantly differ from normal for all factors reported here (Doane & Seward, 2011; Razali & Wah, 2011).

Treatment group effects. The researcher used mixed methods ANOVA with repeated measures to determine if there were any differences in CA scores among the three treatment groups which were the basic game, video and control groups.

The overall mean of the first papers submitted by the basic game group were the highest (M =11.61, SD = 2.11, N = 16) of all the treatment group means. However, again the mean overall CA scores were lower for the student's second papers (M =10.81, SD = 2.12, N = 16). Likewise, the mean overall CA scores for first papers of both the video group (M =11.44, SD = 3.06, N = 13) and the control group (M =11.25, SD = 1.97, N = 22) were both higher than the mean overall CA scores of their second papers as can be seen in Table 5. As illustrated in Table 5, there were no significant differences among the three main treatment groups (basic game, video or control) on either the first or second student analysis papers.

Table 4

Univariate mixed ANOVA Test for treatment group effects

Group	Paper	Overall CA Score	Standard Deviation	N	IRR (α)
Basic Game	1	11.61	2.11	16	.760
	2	10.81	2.12	16	.652*
Video	1	11.44	3.06	13	.760
	2	9.65	2.41	13	.652*
Control	1	11.25	1.97	22	.760
	2	10.08	1.94	22	.652*

Effect of paper order. The researcher also used mixed methods ANOVA with repeated measures to determine if the order in which the papers were analyzed had any effects on the overall CA scores or the scores on the individual CA elements. A total of twenty-five students had read and analyzed Paper A first while the other twenty-six had first read and analyzed Paper B. The group that had analyzed Paper A first had slightly higher CA scores on both papers, but no significant differences were found in the overall CA scores between these two groups. Therefore, the order in which the papers were analyzed (i.e. Paper A first or Paper B first) had no significant effects on student CA scores as illustrated in Table 5.

Table 5

Univariate mixed ANOVA test for paper order effects

Group	Paper	Overall CA Score	Standard Deviation	N
Paper A first	1	11.80	2.10	25
	2	10.27	1.94	25
Paper B first	1	11.03	2.44	26
	2	10.13	2.33	26

Effect of gender. A total of twenty-three men and twenty-eight women participated in the study. The women had slightly higher scores on the first paper and the men had slightly higher scores on the second paper. Also, the difference between the first and second paper for the women (1.4) was greater than for the men (0.98). However, none of these differences were statistically significant.

Table 6

Univariate mixed ANOVA test for gender effects

Group	Paper	Overall CA Score	Standard Deviation	N
Men	1	11.25	2.30	23
	2	10.27	1.90	23
Women	1	11.54	2.32	28
	2	10.14	2.33	28

Intervention Logs

An electronic activity log sheet was sent to all of the study participants in the three original treatment groups. This activity log was also posted on the introductory page of each web-based intervention where students could easily choose to download it. All of the students were asked to use the log to track their usage of the educational intervention and to provide feedback on their assigned intervention.

Only three students returned intervention log sheets to the researcher and all of them had accessed the basic game. No one from the video group submitted a log. As reported in Table 7, all of the students had used either a computer or a laptop to access their assigned intervention. Each submitted log tracked only one access session as shown in Table 7. Two of these logs included comments that the game had been "simple and informative" and "helped somewhat." One student stated that she had previewed the game once and then taken notes on it to "absorb more information," because she had expected more advice to be included. It should also be noted that one of the game logs tracked a single access that had been made three days after the second paper deadline and

Table 7

Intervention log entries

Group	Time Accessed	Type of Device	Times played/watched	Total Time Spent
Game	8:05 p.m.	Laptop	2 times	10 minutes
	2:48 p.m.	Laptop	2 times	9 minutes
	8:03 p.m.	Computer	1 time	3 minutes

so would have been too late to impact the CA skills used to complete the scientific paper analysis assignment.

The LMS used for the interventions (Canvas©) tracked not only total time accessed, but also total number of page views. This turned out not to be helpful in comparing the online intervention options used in this study because watching the entire video counted as only one page view while each slide of the game was tracked as a separate page view.

Study post-test

Students were asked to complete post-tests after the second paper deadline. Only half of the study participants (33 or 48%) had responded by the end of the fall semester. The results of the CA post-test were compared to those of the pre-test for all students who completed both. The CA post-test also asked the student to share their opinions of the online interventions and how helpful they were for completing the scientific study analysis assignment. Students were also asked for their opinions of the CA study and to share their ideas for improving future educational studies of instructional technology.

To encourage more study participants to complete the CA study post-tests, additional drawings for three more $10 bookstore gift certificate were held just prior to the start of the new semester in January, 2015 to encourage participants to complete the CA study post-test. Only one additional person completed the pre-test for a final total of 34 respondents which is 39% of the 69 students who signed consent forms to participate in the study. The researcher only held one more drawing with all of the names of the students who did complete post-tests and delivered the $10 gift certificate early in the spring 2015 semester.

Post-test results: When asked what the two most important factors for determining the strength of a new scientific study were, the post-test respondents gave replies that were very similar to those given during the pre-test. The more frequently a particular word was used in the replies, the larger that word appears in the following word cloud compilation (see Figure 8). It therefore appears that the students felt that Results, Methods, References and Validity were the most important factors. Since this was an open-ended, exploratory question, these terms could then be used in subsequent surveys to provide a more definitive answer to this question.

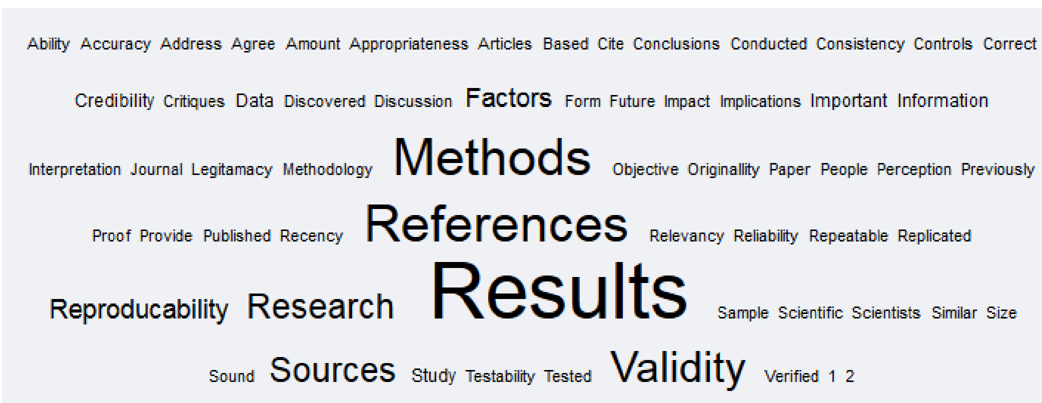


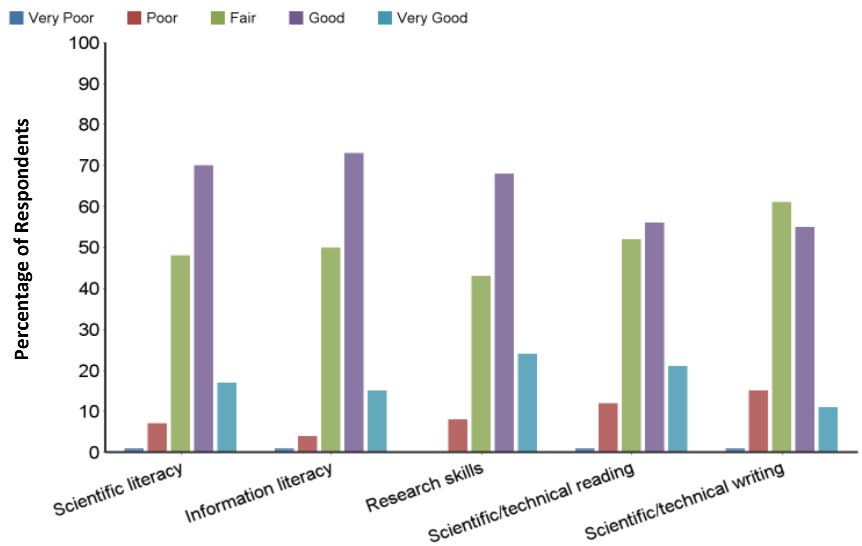
Figure 8: Factors students thought were most important for determining the strength of a new scientific study

The percentage of respondents who correctly identified the reference section as the part of a paper that contains the sources used by the researchers when planning a study rose from 90% in the pre-test to 97% in the post-test. Similarly, the percentage of post-survey respondents who answered that all of the sources used in preparing a scientific research analysis paper, including the article being analyzed, needed to be cited was 97% compared to only 88.6% of the pre-test respondents.

Student ranking of CA skills

When asked to rank their own CA skills and on a scale from 1(poor) to 5 (excellent), most of the students from the pre-test had ranked themselves as either fair or

A) PRE-TEST:



POST-TEST:

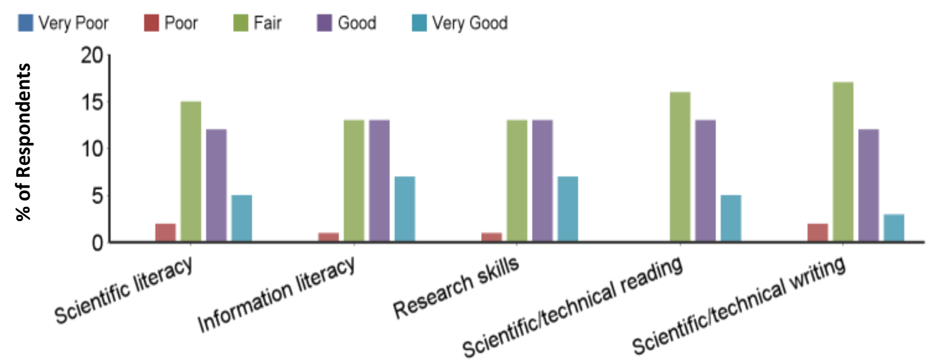


Figure 9: Comparison of pre-test and post-test self-assessment of student CA skills when rated on a scale of 1 (poor) to 5 (excellent)

good as seen in part A of figure 9 and in figure 7 on page 77. This was true of the post-test respondents as well, although the proportion of "good" rankings decreased in the first three categories as illustrated in part B of figure 9. Table 8 compares the means of all of these CA skill indicators. While the student rankings for all but scientific literacy were higher for the post-test than they were for the pre-test, none of these differences were statistically significant.

Table 8

Mean self-assessments of CA skills on a Likert scale from 1 (poor) to 5 (excellent).

CA skill	Mean ranking	Standard Deviation	N
Scientific Literacy			
Pre-test	3.66	0.78	143
Post-test	3.59	0.82	34
Information Literacy			
Pre-test	3.68	0.73	143
Post-test	3.76	0.82	34
Research Skills			
Pre-test	3.76	0.80	143
Post-test	3.76	0.82	34
Scientific Reading			
Pre-test	3.59	0.87	142
Post-test	3.68	0.73	34
Scientific Writing			
Pre-test	3.42	0.81	143
Post-test	3.47	0.75	34

The students were also asked to rate key aspects of the CA study using the same Likert scale of 1 (poor) to 5 (excellent). The results of the study rankings by the students are shown graphically in Figure 10 and quantitatively in Table 9.

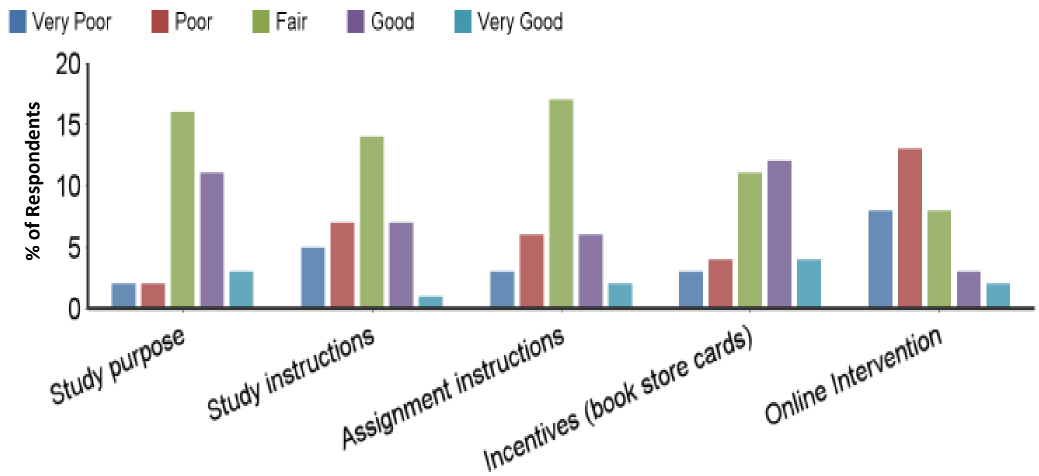


Figure 10: Student ranking of CA study characteristics using a Likert scale of 1 (poor) to 5 (excellent).

Table 9

Student rankings of CA study characteristics on a scale from 1 (poor) to 5 (excellent)

Study Characteristic	Mean Ranking	Standard Deviation	N
Study Purpose	3.32	0.94	34
Study Instructions	2.76	1.05	34
Assignment Instructions	2.94	0.98	34
Incentives (Gift Cards)	3.29	1.12	34
Online Interventions	2.35	1.12	34

Note that the highest mean student ranking of 3.32 (fair) was for the purpose of the study. The incentives (drawings for bookstore incentives) were ranked 3.29 which is also considered to be fair and was the second highest randking. However, the means of the other three study characteristics all fell into the poor category with the online interventions receiving a 2.35 which was the lowest of all of the mean rankings.

The post-test ended with a lot of open-ended questions, the first asking if the students thought that their assigned intervention had helped them with the scientific paper analysis. Most of the responses, twenty-four out of a total of thirty-two (75%) said no. Additional information added to negative responses indicated that the students thought their intervention was too short ("just a quiz"), off subject ("I didn't understand the point of the game"), too easy ("it was a quiz that I knew the answers to" or "common knowledge"), frustrating or in general "had nothing to do with the assignment." Obviously most of the students would prefer explicit instruction over needing to discover the connection and transfer the knowledge to the task at hand, which was not unexpected. Four students (12.5%) said yes, their intervention had helped because "it gave me an idea of what the professors were looking for" and another student said that "it helped me write my paper." One student said "Yes, because it made me refocus my perspective on how to actually critique others' work rather than simply stating what was found. " A handful of students responded without directly answering the question. One of these students wrote "I did not know what the intervention was until just now" while two others mentioned glitches and one person talked about not knowing how to access their intervention.

When asked if they had viewed other study interventions, only twenty-one students answered (52%) but the answer was a resounding "no." However, to the follow-up question about which ones they had reviewed, one student answered that he had spent several hours on each one. While the other nine students wrote in not applicable or n/a.

When asked which study intervention they like the most, the video was mentioned by three of the eleven people who answered this question, while the game was mentioned by one. One of the students answered "LU professors," which was counted as a vote for the video. One person said "the one about the autism" apparently confusing the assigned papers with the interventions. The other six people responded with "no" or "n/a."

Only ten people responded to the question about which study intervention helped them the most with the scientific article analysis. Two of these mentioned the video while the tutorial and the game were both mentioned by one person. This means that each of the study interventions had helped at least one person.

The second to the last question asked the students if they would participate in this type of study again. Fifteen students out of the thirty-two who answered this question (47%) said yes, thirteen (41%) said no and the remaining four people (12%) said maybe. Sample comments as to why they would or would not participate again included:

> "Yes, it did not require much time and it ran parallel with required assignments."
>
> "No because I did not know about the first one before I was expected to complete the second which was not conducive to making improvements."
>
> "Yes, because I think it is important to realize that most students do not know how to write these sorts of papers."
>
> "I don't think I am in a position to answer that un-biasedly, as an overworked college student, I want as little work as possible, but if it was less 'Do more

> homework!' and more like an actual 'video teaching experience" where you get credit for watching the videos/going through these 'study interventions' and writing a summary of them/report on them, that would be a very helpful learning experience."

The students were very willing to share ideas for improving the study, like the last quote listed above, even when they weren't directly asked for them. Since the rest of the post-test questions asked specifically for ways to improve the study interventions and the study itself, they will be discussed in more detail in the last chapter of this dissertation; the discussion section.

Post –test summary

The students liked the purpose of the study and the opportunity to improve what many of them already understand to be important skills. However, they were disappointed by the interventions themselves, angry that they would not know their grade on the first paper prior to writing the second and felt frustrated by many aspects of the study from its timing ("when I already had four exams that week") to the glitches to feeling that were "put into a small group, then forgotten" presumably by what they perceived as a lack of communication.

The researcher thought she had clearly provided all the information they needed, but it turned out that it wasn't where they wanted it or presented in the way they wanted to receive it. Many students mentioned placing all of the information in the assignment folder or at least within CourseSite instead of having it "buried in e-mail."

Some of the problems created by using an outside LMS that sent e-mail invitations to students, that they then had to recognize as part of this study, were addressed in January, 2015 when CourseSite upgraded to the newest release of Moodle, the backbone of CourseSite. This latest version of Moodle has an option called grouping which allows the instructor to add course content (assignments, resources) that is only visible to certain students. This function was offered as an experimental feature in the last release of Moodle and the university's Library and Technology Services team made the conscious decision to not use it until it had been fully tested by others. As the researcher was often told in the last few months, this study was just a semester too early to take full advantage of this powerful feature.

In summary, the results of this study provided no clear answers about the impact of content interactivity within online learning resources. However, the results do emphasize the importance of taking the time to deeply involve members of the target audience in the development of any new web-based resources. The student comments received via the pre-test, post-test, logs and e-mails also revealed just how important it is to try to emulate the kinds of interfaces these students are used to dealing with on a daily basis when trying to entice them to use a newly designed online educational intervention.

Chapter 5
Discussion

Scientists need to develop high levels of critical analysis (CA) skills including how to critically read new scientific information and how to communicate, especially in writing. These skills are vital to both sharing research findings and obtaining funds to be able to continue their research. This study explored alternative ways to provide online educational resources for developing these skills in early-year college students interested in a transformative field, genetics. It also refined a tool for measuring development of CA skills in the form of a rubric which could provide the framework for scaffolding science education in higher education and enhance the 21st Century skills needed to tackle urgent global problems such as climate change, food production and sustainable energy.

Many of the bright, aspiring young scientists who participated in this study were at once challenged and thrilled at the prospect of learning scientific writing. Yet, as one of the students who completed the post-test commented they were in the end "really, really disappointed that my study intervention did not teach me how to write a better paper." We need to find ways to better meet their needs and involve them more in shaping these skills so vital to their future careers and to the survival of our planet.

The challenge for instructional designers is to design learning environments where students of many different abilities and needs can all find what they need to take their own communication skills to the next level. Sometimes it is valuable to review more basic material so that students can correct any misperceptions that can result from a shallow overview of material (Khan, 2012) and instead support the foundational understanding needed to support additional learning (Jonassen & Land, 2012). For example, in this study, almost all of the students addressed the CA elements in the order in which they appeared in the assignment. The first CA element dealt with the purpose of the research study and the point of view of the researchers. Many students were confused by what was meant by point of view. Instead of trying to think of it in context with the other CA elements or even ask themselves what point of view they might have about the field of genetics, they instead performed a cursory web search for the term. The first entry to commonly appear in such a search yields the grammatical use of point of view as "the perspective from which a speaker or writer recounts a narrative or presents information" (Wikipedia, 2015). Since this definition is related to writing in general, they assumed that this was what they should include in their scientific analysis papers. These students totally neglected to notice another commonly provided web search result for this term in which point of view refers to "the angle of considering things, which shows us the opinion, or feelings of the individuals involved in a situation," such as the study authors, which could possibly bias scientific studies and their subsequent findings (http://literarydevices.net/point-of-view/). These students may continue to make this mistake thinking that a professional analysis of the credibility of a scientific paper needs to check that the author has correctly used an objective voice throughout. This leads to a focus on the mechanics of writing over the essence of the scientific methods and meaning that is not as likely to lead to enhanced *critical analysis skills* nor enhance critical thinking itself (Poniastowski, 2012; Terry, 2012)

Conclusions

The students were more likely to access an online educational game that than they were to access either a video or a tutorial. In addition, those who accessed the game played it at least once, with most of them interacting with it more than their cohorts interacted with either the video or the tutorial. In fact, not every student who visited the introductory page of the video or the tutorial chose to interact with these educational interventions at all. Moreover, a third of the 69 students who originally signed consent forms to participate in the study did not access the educational interventions available to them even once. Enticing college students to access all the online resources available to them in the multimedia-rich world in which they are immersed can be challenging (Anderson, 2003; Murray et al., 2013). Even the simplest aspects of accessing a lesson need to evaluated from their perspective. For example, had the end of the pre-test automatically taken students to the next step or if the verification code would have started flashing, as they often do on commercial websites, it would have prevented the small, but vocal minority of students who struggled with the first steps in the assignment website from becoming so frustrated.

Even when accessed, the three online interactions did not produce the intended impact. Whether this was because they were too short, not appealing enough or not perceived by the students to be related enough to the assignment (Murray et al., 2013) is difficult to tell given the overall circumstances. Most likely the students felt so overwhelmed by midterms during the week they were expected to review online interventions and write their second paper analysis, that they did not put forth their best efforts on writing their second scientific study analysis paper. The fact that some people plagiarized, when none had on the first paper, is further indication that the students were so pressed for time that they were willing to try anything just to not lose the points.

However, it could also be that the students were so put off by not receiving feedback on their first paper, that they did not approach the second paper with the same level of dedication and rigor. In other words, the lack of feedback had sapped their motivation to even try to learn what they had thought they would be learning along with their desire to do the best they could (Tollefson, 200). This is even more likely when students are performing tasks, such as this assignment, that they perceive to be difficult (Bandura, 1993). Educators need to build up, not tear down motivation to learn and perform at higher levels in order to help students build the skill they will need for life in the 21st century (Bandura, 1993; p21, 2010). Future studies should be more mindful of the need to set better student expectations from the beginning (Tollefson, 2000).

Clearly for this study, the problem was that the instruction the students received was not what they had anticipated which may have led to disappointment, frustration and a lack of motivation to apply their best efforts to the task of completing a second analysis paper assignment. These feelings could have been magnified by the stress of having the second paper due during mid-term week. Designing better lessons with liberal amounts of student input is one way to address this. The students had many positive suggestions, which will be discussed as recommendations for future studies.

Other important lessons that can be gleaned from the results of this study are that when dealing with innovative education, there is only one chance to make a good first impression. Confusion at the start of a study is not going to create the kind of positive, or at least neutral environment that will keep the subjects open to the new ideas your are testing on them. Taking the time to establish realistic expectations and ensure that the subjects understand and agree to all of the aspects of a study are important in maintaining the kind of openness needed when people are trying new and exciting behaviors, learning objects or ways of thinking (Dick, Carey & Carey, 2009; Wiggins & McTighe, 2005).

The low level of internal validity and IRR for some of the individual CA elements makes it impossible to fully trust any differences found in these key areas. However, it was encouraging to see improvement in the student's ability to analyze and discuss the credibility of the references used in the analyzed study and recognize the implications of the studies.

In this study, the researcher analyzed the need for CA skills more than she analyzed her target audience. She designed the interventions based on the literature on science writing and on what had been warmly received when she presented her prototypes to other educators at national conferences. In hindsight, she had neglected to fully analyze her target audience. She had solicited feedback from a handful of college students who were either writing tutors or related to the researcher, but may not necessarily have been representative of the majority of college science students.

Implications

This study revealed that early-year college science students do understand the importance of critical thinking in their future careers and are interested in further building their critical analysis skills such as science writing along with the science processing skills they are honing in the lab. As one student phrased it**,** he/she liked the study "because it is directed toward making science students better critical thinkers and because scientific writing is a valuable skill to have especially when performing any type of research."

Students have varying degrees of CA skills. Some clearly felt that the information included in the online interventions was not challenging enough, while others appreciated the basic information. This indicates a need for interventions that offer multiple levels of instruction through scaffolding or branching.

Some of the study participants found the online educational interventions designed for this study somewhat useful, but many commented that they already knew the information contained in them. Most of the participating students indicated that they wanted more explicit instruction on exactly how to structure an analysis paper. There is a fine line between helping a student learn and providing them with all the answers. It is especially important to the development of critical thinking that educators allow an adequate amount of challenge and ambiguity that can foster the kind of problem-solving needed for deep learning (Jonassen et al., 2008). This study vividly demonstrated what happens when instructional designers develop specialized e-learning without a thorough assessment of needs, interests and educational preferences of the target audience (Allen et al., 2014). Their input would no doubt have led to a different time table as well.

The researcher intends to re-work the intervention that did not work and illicit the reaction of the student fellows of the Writing Across the Curriculum program so that the new and improved interventions can be used by all to improve science writing skill instruction at this university.

As time and resources allow, she will hold focus group meetings with high school students and other potential users of these resources so that more effective educational programs can be continually developed to meet the future needs of the scientific workforce of the 21st Century skills (p21, 2011).

Limitations

It is difficult to add an assignment such as this to an already packed academic and social schedule. In their efforts to avoid social conflicts, the researcher and instructor inadvertently created a greater academic burden for the students. The second wave of midterms appears to have decreased the amount of time that the students had both to access their study interventions and to read their second primary research article and prepare their second analysis paper.

It is difficult to find qualified raters for prolonged rating sessions. The researcher had several potential raters that had been willing to work with her earlier, but only two who were available for an extended period of time after the final paper deadline in November, 2014. Unfortunately, both of the raters who did commit to this study had full-time jobs that already took more than 40 hours a week of their time along with other significant commitments that prevented them from meeting together with the researcher for more than four-five hours per week.

A lack of good-quality, student papers for the validity protocol was a constant challenge. The researcher tried to use fresh papers as much as possible, but only about forty papers from the pilot study were useful for determining internal and inter-rater reliability. The raters soon recognized ones that had been used before, which may have impacted the ability to achieve adequate IRR in this study.

In addition, both of the volunteer raters had little to no background in the natural sciences and the papers were much more challenging to understand both for the raters and the students themselves than were the papers that the students analyzed in the pilot study. The researcher therefore decided to write a guide to the scientific information within each of the assigned papers to help clarify the content, findings and implications of the assigned papers. Before she gave the guides to the raters, the researcher asked a molecular biology professor to review them for accuracy and completeness. After reviewing both the assigned papers and the guides, this colleague felt strongly that the papers were too complex for raters who were not familiar with cellular reproduction, genetics and other foundational biology concepts. She suggested that graduate students in one of these fields, college science teachers or others with equally strong science backgrounds might rate the papers very differently when compared to the two current raters. Perhaps some of the variability in the ratings stemmed from a lack of full understanding of the underlying scientific information.

The researcher contacted several people who were not able to join in the rating process until she contacted an acquaintance who had taught high school biology for over thirty years and served as an officer in the National Association of Biology Teachers. This acquaintance even had served as a rater in previous educational studies. She had not been available to serve as a rater in the past, but this time, she had time to work with the researcher on a short-term basis to determine if her science expertise was indeed significant to the overall rating process.

The researcher trained this new rater and had her complete several sets of qualifying papers that she compared with the same from the two raters who had been working together for months. The new rater's scores correlated adequately with the other raters and so all of the raters met around the same time that the original raters were less than 20 papers away from reaching the target sample size of 51 subjects. After correlating again with the other raters using the pilot study papers, the researcher gave this new rater 19 experimental papers to score. Unfortunately, the scores on these new papers did not correlate as well with the other raters, indicating that future studies should attempt to secure commitments from science savvy raters from the outset of the study.

Another limitation was the short instructional time for each treatment. The researcher deliberately attempted to match the length of the intervention to the average adult attention span. This may have led to a more superficial coverage of the topics which could be addressed by using these interventions as introductory pieces which would branch off to more detailed information about the concepts encompassed by the CA skill rubric and how to apply them when analyzing scientific research or in science writing. The researcher already has plans to make two different follow-up videos using parts of the professor videos that were not used in making the expert video for this study.

A final limitation was the relatively small number of subjects within each treatment cell, which could be addressed by offering this type of intervention either in larger course sections or in more science courses, which would simultaneously enhance the generalizability of the results. However, scaling up the size of the potential pool would simultaneously increase the workload for the raters as well as the researcher. Automated essay scoring (AES) mechanisms, such as the ones used in massively open online courses (MOOCs) or high-stakes testing such as college entrance exams could be used at least to assign student grades. Although any type of feedback on pre-treatment papers would still have to be withheld until the source of the experimental data (in this study, the second paper) has been gathered. Perhaps the CA rubric could be integrated into a specialized AES system for enhancing scientific writing skills.

Recommendations for further study

This study validated that at least some science students realized the importance of CA skills for their future careers and are very interested in learning how to read the primary literature and prepare written analysis papers. This was evident in many of the comments made in the pre-test, post-test, intervention logs and throughout the semester via e-mails to the researcher.

Some students indicated in the post-test results that their online intervention had been at least somewhat helpful. Yet, the tone of most of the post-test feedback was that the instruction had been too basic and not practical enough. This dichotomy could be have been handled by creating branched interventions that allow students to skip over what they already understand and more easily get to new information that they might perceived to be more used. This could lead to a specialized curriculum designed to provide deep learning that is responsive to varying and changing student needs and interests. The results of this study underscore that the importance fully understanding these desired outcomes prior to developing educational technology. Anyone mounting similar studies needs to administer pre-study surveys of student needs and hold open meetings with subjects similar to the target group in order to determine exactly what the students may know, what they would like to learn and exactly how they would most like to receive the information (Jonassen, 2008). Ideally, story boards and/or simple prototypical lessons could be shared with student representatives to gain their feedback prior to settling on the final instructional design (Allen et al., 2014; Dick et al., 2009). Students who are similar to the intended target audience could be included throughout the development process to enhance the likelihood that the final lesson would be well-received and effective from the student's point of view (Allen et al., 2014). This process is similar to the Sequential Approximation Model (SAM) of technology-based instructional design in which simple prototypes of new ideas are created and shared with all stakeholders, ideally during brainstorming and planning meetings or shortly afterwards (Allen & Sites, 2012). Students could even be encouraged to design sample online lessons themselves (Anderson, 2003).

In addition, all of this planning and creating should be completed at least a few months before the actual implementation of the new learning modules to ensure that technical problems are minimized (Allen et al., 2014). It can be a challenge to choose design elements and ideas that have had adequate testing but do not appear to be out-of-date by the time they are actually fully launched.

The students who participated in this study shared many great ideas for future work in this important area including creating an interactive sample paper, providing lists what to do and not to do when analyzing or writing scientific papers and providing more detailed instructional videos including videos of doctoral candidates explaining the techniques they use for scientific writing/communication. This last idea is an extension of the value of a novice learner who understands how to do something but still remembers how hard it was to understand at first (Watkins & Mazur, 2013).

A final recommendation from a professor at the university was to greatly shorten the time frame between writing the papers so that the interventions can be tested without extraneous variables such as the ones that arose in this study because of the shift in the timeline. Some of the standardized tests already require two essays within a short time frame and it might be possible to mimic this in the science curriculum as well. Perhaps, students could be asked to write a brief analysis followed immediately by a well-tested intervention and an opportunity to write another analysis paper. The reward for them could be that they will get assigned the highest score, but the bigger reward would be the more effective educational interventions that would result from such a controlled study.

References

Achieve, Inc. (2013). *Next Generation Science Standards*. Washington, DC: Achieve, Inc. Retrieved from http://www.nextgenscience.org/

Achison, C., & Lee, A. (2006). Research writing: Problems and pedagogies. *Teaching in Higher Education, 11,* 265-278.

Adams, C. C., & Columa, H. L. (2014). Classroom Response Systems: Effects on the Critical Thinking Skills of Students in Introductory Science Courses. *School Science and Mathematics, 114*(8), 362-379.

Allen, I. E., & Seaman, J. (2010). *Class differences: Online education in the United States, 2010.* Babson College, MA: Sloan Consortium. Available from http://sloanconsortium.org/publications/survey/class_differences

Allen, I. E., & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States*. Babson College, MA: Sloan Consortium. Available from http://sloanconsortium.org/publications/survey/changing_course_2012

Allen, M., & Sites, R. (2012). *Leaving ADDIE for SAM*. Alexandria, VA: ASTD Press.

Allen, M. (2014). Join the e-learning revolution. *Training, 51*(2), 52.

Allen, M., Dirksen, J., Quinn, C., & Thalheimer, W. (2014). Serious e-learning manifesto. Retrieved 4/24/2014 from http://elearningmanifesto.org/read-the-manifesto/

Albitz, R. S. (2007). The what and who of information literacy in higher education. *Libraries and the Academy, 7*(1), 97-109.

American Association for the Advancement of Science (AAAS). (1989). *Science for all Americans: A project 2061 report on literacy goals in science, mathematics and technology*, Washington, DC: American Association for the Advancement of Science.

American Association for the Advancement of Science (AAAS). (1993). *Benchmarks for science literacy.* New York, NY: Oxford University Press.

American Association for the Advancement of Science (AAAS). (2011). Vision and change in undergraduate biology education: A call to action. Retrieved from http://www.visionandchange.org

Amiel, T., & Orey, M. (2007). Do you have the time? Investigating online classroom workload. *Journal of Educational Technology Systems, 35*(1), 31-43. doi: 10.2190/CU8Q-8678-4W03-3587

Anderson, L. W. & Krathwohl, D.R. (Eds) (2001). A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of education. New York, NY: Longman.

Anderson, T. (2003a). Getting the mix right again: An updated and theoretical rationale for interaction. *International Review of Research in Open and Distance Learning, 4*(2). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/149/230

Anderson, T. (2004). Toward a theory of online learning. In T. Anderson & F. Elloumi (Eds.), *Theory and practice of online learning.* pp. 33-60. Athabasca, AB, Canada: Athabasca University Press. Retrieved from http://cde.athabascau.ca/online_book/ch2.html

Annetta, L. A., & Shymansky, J. A. (2006). Investigating science learning for rural elementary school teachers in a professional development project through three distance-education strategies. *Journal of Research in Science Teaching, 43,* 1019-1039. doi: 10.1002/tea.20106

Appana, S. (2008). A review of benefits and limitations of online learning in the context of the student, the instructor and the tenured faculty. *International Journal of ELearning, 7*(1), 5-22.

Association of College and Research Libraries (ACRL). (2000). *Information Literacy Competency Standards for Higher Education.* Chicago, IL: American Library Association. Retrieved from www.ala.org/acrl/standards/informationliteracycompetency

Astleitner, H. (2002). Teaching critical thinking online. *Journal of Instructional Psychology, 29(*2), 53-76.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4,* pp. 829-837.

Balgopal, M., & Wallace, A. (2013). Writing-to-Learn, Writing-to-Communicate, & Scientific Literacy. *American Biology Teacher*, *75*(3), 170-175.

Baker, W. P., Barstack, R., Clark, D., Hull, E., Goodman, B. Kook, J., Kraft, K., Ramakrishna, P., Roberts, E., Shaw, J., Weaver, D., & Lang, M. (2008). Writing-to-learn in the inquiry-science classroom: Effective strategies from middle school science and writing teachers. *The Clearing House, 81*(3), 105-108.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychology, 28*, 117-148.

Banko, W., Grant, M. L., Jabot, M. E., McCormack, A. J. & O'Brien, T. (2013). *Science for the next generation.* Arlington, VA: NSTA (National Science Teachers Association) Press.

Bichsel, J, (2013). *The state of e-learning in higher education: An eye toward growth and increased access* (Research Report), Louisville, CO: EDUCAUSE Center for Analysis and Research. Retrieved from http://www.educause.edu/ecar

Bloom, B. S. & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners. Handbook1: Cognitive domain. New York, NY: Longman.

Boettcher, J. (2013). Ten best practices for teaching online. Retrieved from http://www.designingforlearning.info/services/writing/ecoach/tenbest.html

Broussard, M. S. (2012). Digital games in academic libraries: a review of games and suggested best practices. *Reference Services Review*, *40*(1), 75-89. doi:10.1108/00907321211203649

Brunce, D. M., Flens, E. A., & Neiles, K. Y. (2010). How long can students pay attention in class? A study of student attention decline using clickers. *Journal of Chemical Education, 87*(12), 1438 – 1443. doi: 10.1021/ed100409p

Bybee, W. R. (1997). Towards an understanding of scientific literacy. In W. Gräber & C. Bolte. (Eds.). *Scientific literacy: An international symposium* (p. 37-68). Institut für die Pädagogik der Naturwisenschaften (IPN): Kiel, Germany.

Bybee, W. R., Taylor, J. A., Gardner, A., Scotter, P. V., Powell, J. C., Westbrook, A., & Landes, N. (2006). *The BSCS 5E instructional model: origins, effectiveness, and applications.* Colorado Springs, CO: BSCS.

Çavdar, G., & Doe, S. (2012). Learning through writing: Teaching critical thinking skills in writing assignments. *Ps: Political Science and Politics, 45,* 2, 298-306. doi: 10.1017/S1049096511002137

Clarke, R. E. (1994). Media will never influence learning. *Review of Educational Research, 53*(4), 445-459.

Clary, R. M. & Wandersee, J. H. (2012). Mandatory climate change discussions in online classrooms: Promoting students' climate literacy and understanding of the nature of science. *Journal of College Science Teaching, 41*(5), 70-79.

Coates, H. (2007). A model of online and general campus-based student engagement. *Assessment & Evaluation in Higher Education, 32*(2), 121-141. doi: 10.1080/02602930600801878

Coil, D., Wenderoth, M. P., Cunningham, M. & Dirks, C. (2010). Teaching the Process of Science: Faculty Perceptions and an Effective Methodolody. *CBE – Life Sciences Education, 9*(4), 524-535. doi: 10.1187/cbe.10-01-0005

Collins, A., & Halverson, R. (2009). *Rethinking education in the age of technology*. New York, NY: Teachers College Press.

Cook, M. P. (2006). Visual representation in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education, 90*(6), 1073-1091. doi: 10.1002/sce.20164

Corradi, H. R. (2012). Student centered education: A sample practical report to facilitate writing in the scientific style, *Biochemistry and Molecular Biology Education, 40*(1), 65-67. doi:10.1002/bmb.20571

Costa, A. L., & Kallick, B. (2000). *Assessing and reporting on habits of mind.* Alexandria, VA: Association for Supervision and Curriculum Development.

Dangel, H. L., & Wang, C. X. (2008). Student response systems in higher education: Moving beyond linear teaching and surface learning. *Journal of Educational TechnologyDevelopment and Exchange, 1*(1), 93-104. Retrieved 5/8/2013 from: http://www.sicet.org/journals/jetde/jetde08/paper08.pdf

Davis, B., Carmean, C., & Wagner, E. D. (2009). *The evolution of the LMS: From management to learning.*Santa Rosa, CA: The eLearning Guild. Retrieved from: http://www.educationtrainingsolutions.com/wp-content/uploads/2009/10/ evolution-of-LMS.pdf

Dewey, J. (1938). *Experience and Education*. Toronto: Collier-MacMillan Canada Ltd.

Dick, W., Carey, L., and Carey, J. O. (2009). *The systematic design of instruction,* (7th Ed.). Upper Saddle River, NJ: Pearson.

Doane, D. P., & Seward, L. E. (2011). Measuring Skewness. *Journal of Statistical Education, 19*(2), 1-18.

Dunlap, J. C., Sobel, D., & Sands, D. I. (2007). Designing for Deep and Meaningful Student-to-Content Interactions. *Techtrends: Linking Research & Practice To Improve Learning*, *51*(4), 20-31. doi:10.1007/s11528-007-0052-6

Edelson, D., & Reiser, R. A. (2006). Making authentic practices accessible to learners: Design challenges and strategies. In R. K. Sawyer (Ed.). *The Cambridge Handbook of the Learning Sciences* (p. 335-354). Cambridge, MA: Cambridge University Press.

Ediger, M. (2012). Writing in the science curriculum. *Education, 133*(1), 35-38.

Elder, L., & Paul, R. (2007). The miniature guide to taking charge of the human mind: Thinking, feeling, wanting. 3rd Ed. Dillon Beach, CA: Foundation for Critical Thinking.

Ennis, R. H. (1991). Critical thinking: A streamlined conception. *Teaching Philosophy, 14*(1), 5-24.

Ennis, R. H. (2013). The nature of critical thinking: Outlines of critical thinking dispositions and abilities. Retrieved from http://www.criticalthinking.net/longdefinition.html

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction.* Millbrae, CA: The California Academy Press. Available at http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet ?accno=ED3154232

Facione, P. A., Facione, N. C., Giancarlo, C. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informational Logic, 20*(1), 61-84. doi:

Fan, X., Geelan, D. (2013). Enhancing students' scientific literacy using interactive simulations: A critical literature review. *Journal of Mathematics and Science Teaching, 32*(2), 125-171.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics, 4th Edition.* Thousand Oaks, CA: SAGE Publications Inc.

Fish, W. W., & Wickersham, L.E. (2009). Best practices for online instructors: Reminders. *The Quarterly Review of Distance Educational Leadership, 10,* 279-284.

Fisher, S. L., Wasserman, M. E., & Orvis, K. A. (2010). Trainee reactions to learner control: An important link in the e-learning equation. *International Journal of Training and Development, 14*(3), 198-208.

Flynn, A. (2011). Developing problem-solving skills through retrosynthetic analysis and clickers in organic chemistry. *Journal of Chemical Education, 88*, 1496-1500. doi: 10.1021/ed200143k

Fournier, H., Kop, R., & Durand, G. (2014). Challenges to research in MOOCs, *MERLOT Journal of Online Learning and Teaching, 10*(1), 1-15.

Fosmire, M. (2010). Calibrated peer review: A new tool for integrating information literacy skills in writing-intensive large classroom settings. *Libraries and the Academy, 10*(2), 147-163. doi: 10.1353/pla.0.0099

Fraser, J. M., Timan, A. L., Miller, K., Dowd, J. E., Tucker, L., & Mazur, E. (2014). Teaching and physics education research: Bridging the gap. *Reports on Progress in Physics, 77*(3). doi: 10.1088/0034-4885/77/3/032401

Fry, S. W., & Villagomez, A. (2012). Writing to learn: Benefits and limitations. *College Teaching, 60*(4), 170-175. doi: 10.1080/87567555.2012.697081

Fulka, H., Mrazek, M., Tepla, O. & Fulka, Jr., J. (2004). DNA methylation pattern in human zygotes and developing embryos. *Reproduction, 128.* 703-708. doi: 10.1530/rep.1.00217

Fulton, K. (2012). Upside down and inside out: Flip your classroom to improve student learning. *Learning & Leading with Technology*, 39(8), 12-17.

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical thinking in a text-based environment: Computer conferencing in higher education. *Internet and Higher Education, 2*(2-3), 87-105.

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence and computer conferencing in distance education. *American Journal of Distance Education, 15*(1), 7-23.

Glaser, E. M. (1941). *An experiment in the development of critical thinking.* New York, NY: Teacher's College Press.

Greene, L. (2010). Writing in the lifes sciences: A critical thinking approach. New York: Oxford University Press.

Greenlaw, S. A., & DeLoach, S. B. (2003). Teaching critical thinking with electronic discussion. *The Journal of Economic Education, 34*(1), 36-52.

Greenstein, G. (2013). Writing is thinking: Using writing to teach science. *Astronomy Education Review, 12*(1), doi: 10.3847/AER2012037

Gredler, M. (2009). *Learning and instruction: Theory into practice.* Upper Saddle River, NJ: Pearson.

Gunersel, A. B., Simpson, N. J., Aufderheide, K. J., & Wang, L. (2008). Effectiveness of Calibrated Peer Review™ for improving writing and critical thinking skills in biology undergraduate students. *Journal of the Scholarship of Teaching and Learning, 8*(2), 25-37. Retrieved 4/22/14 from http://www.eric.ed.gov/content delivery/servlet/ERICServlet?accno=EJ854838

Gunn, T. M., Grigg, L. M., Pomahac, G. (2006). Critical thinking and bioethical decision making in the middle school classroom. *International Journal of Learning, 13*(5), 128-135. Retreived 3/29/2013 from http://www.Learning-Journal.com, ISSN 1447-9494

Halverson, K. L., Siegel, M. A., & Freyermuth, S. K. (2010). Non-science majors' critical evaluation of websites in a biotechnology course. *Journal of Science Education Technology, 19*(6), 612-620. doi: 10.1007/s10956-010-9227-6

Hannafin, M. J., Hannafin, K. M., Land, S. M., & Oliver, K. (1997). Grounded practice and the design of constructivist learning environments. *Educational Technology Research and Development*, *45*(3), 101-117. doi: 10.1007/BF02299733

Hannafin, M., Hannafin, K., & Gabbitas, B. (2009). Re-examining cognition during student-centered, web-based learning. *Educational Technology Research and Development*, *57*(6), 767-785. doi: 10.1007/s11423-009-9117-x

Harmon, O. R., Lambrinos, J., & Buffolino, J. (2010). Assessment design and cheating risk in online instruction. *Online Journal Of Distance Learning Administration*, *13*(3). Available at: http://www.westga.edu/~distance/ojdla/Fall133/harmon_lambrinos_buffolino133.html

Holbrook, J., & Rannikmae, M. (2009). The Meaning of Scientific Literacy. *International Journal of Environmental & Science Education*, *4*(3), 275-288.

Hoskins, S. G. (2010). But if it's in the newspaper, doesn't that mean it's true?. Developing critical reading and analysis skills by evaluating newspaper science with CREATE. *The American Biology Teacher, 72*(7), 415-420.

Huffman, D., & Heller, P. (1995). What does the force concept inventory really measure?. *The Physics Teacher, 33*(3), 138-143.

Huerta, D., & McMillan, V. (2004). Reflections on collaborative teaching of science information literacy and science writing: Plans, processes and pratfalls. *Resource Sharing & Information Networks*, *17*(1/2), 19-28. doi:10.1300/J121v17n0103

Inman, C., Wright, V. H., & Hartman, J. A. (2010). Use of Second Life in K-12 and Higher Education: A Review of Research. *Journal of Interactive Online Learning*, *9*(1), 44-63.

International Association for K-12 Online Learning (iNACOL). (2011). *The Online Learning Definitions Project.* Retrieved from: http://www.inacol.org/cms/wp-content/uploads/2013/04/iNACOL_DefinitionsProject.pdf

Jackson, T. R., Draugalis, J. R., Slack, M. K., & Zachry, W. M. (2002). Validation of authentic performance assessment: A process suited for Rasch Modeling. *American Journal of Pharmaceutical Education, 66*, 233-242.

Janick-Buckner, D. (1997). Getting undergraduates to really read and discuss primary literature. *Journal of College Science Teaching, 27*(1)*,* 29-32.

Jonassen, D. H., Carr, C., & Yueh, H. P. (1998). Computers as mindtools for engaging learners in critical thinking. *TechTrends*, *43*(2), 24-32.

Jonassen, D. H. (2008). Instructional design as design problem solving: An iterative process. *Educational Technology, 48*(3), 21-26.

Jonassen, D., Howland, J., Marra, R. M., & Crismond, D. (2008). *Meaningful learning with technology.* Upper Saddle River, NJ: Pearson Education, Inc.

Jonassen, D., & Land, S. (2012). *Theoretical Foundations of Learning Environments.* New York, NY: Routledge. Available for preview at: http://site.ebrary.com/ lib/alltitles/docDetail.action?docID=10534988&ppg=2

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130-144

Khan, S. (2012). *The one world school house.* New York, NY: Twelve, Hachette Book Group.

Kalman, C. S. (2011). Enhancing students' conceptual understanding by engaging science text with reflective writing as a Hermeneutical circle. *Science & Education, 20,* 159-172. doi: 10.1007/s11191-010-9298-z

Kandel, E. R. (2006). In search of memory: The emergence of a new science of mind. New York, NY: W. W. Norton & Company.

Ke, F., & Hoadley, C. (2009). Evaluating Online Learning Communities. *Educational Tech-nology Research and Development*, *57*(4), 487-510. doi: 10.1007/s11423-009-9120-2

Keller, J. M. (2008). First principles of motivation to learn and e^3-learning. *Distance Education, 29*(2), 175-185. doi: 10.1080/01587910802154970

Kennedy, G. (2004). Promoting cognition in multimedia interactivity research. *Journal of Interactive Learning Research, 15*(1), 43-61.

Kousta, S.-T., Vigliocco, G., Del Campo, E., Vinson, D. P., & Andrews,M. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General, 140*, 14–34. doi:10.1037/a0021446

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development, 42*(2), 7-19.

Lapadat, J. C. (2006). Written interaction: A key component in online learning. *Journal of Computer-Mediated Communication, 7*(4), 0. doi: 10.1111/j.1083-6101.2002.tb00158

Lee, J., & Spector, M. (2012). Effects of model-centered instruction on effectiveness, efficiency, and engagement with ill-structured problem solving. *Instructional Science, 40*(3), 537-557. doi: 10.1007/s11251-011-9189-y

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2008). *SPSS for intermediate statistics: Use and interpretation.* (3rd Ed.). New York, NY: Psychology Press, Taylor & Francis Group.

Libarkin, J., & Ording, G. (2012). The utility of writing assignments in undergraduate biology. *CBE-Life Sciences Education, 11*, 39-46. doi: 10.1187/cbe.11-07-0058

Lord, T., & Orkwiszewski, T. (2006). Moving from didactic to inquiry-based instruction in a science laboratory. *American Biology Teacher*, *68*(6), 342-345.

Lynd-Balta, E. (2006). Using literature and innovative assessments to ignite interest and cultivate critical thinking skills in an undergraduate neuroscience course. *CBE-Life Sciences Education, 5,* 167-174. doi: 10.1187/cbe.05-08-0108

Maeder, D. W. (2007). Online learning and assessment solutions. *National Business Education Yearbook, 45,* 102-116.

Malliarakis, C., Satratzemi, M., & Xinogalos, S. (2014). Designing Educational Games for Computer Programming: A holistic Framework. *Electronic Journal Of E-Learning*, *12*(3), 281-298.

Marin, L. M., & Halpern, D. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity, 6*, 1-13. doi: 10.1016/j.tsc.2010.08.002

Marschall, S., & Davis, C. (2012). A conceptual framework for teaching critical reading skills to adult college students. *Adult Learning, 23*(2), 63-68. doi: 10.1177/1045159512444265

Mauldin, R. F. (2011). A novel approach to teaching scientific reasoning to future journalists: An intellectual framework for evaluating press reports about scientific research. *Science Communication 34*(2), 283-291. doi: 10.117/1075547011427976

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

Means, B., Bakia, M., & Murphy, R. (2014). Learning online: What research tells us about whether, when and how. New York, NY: Routledge.

Merrill, M. D., & Drake, L. (1996). Reclaiming instructional design. *Educational Technology*, *36*(5), 5

Meyer, J. H. F., & Land, R. (2005). Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. *Higher Education, 49*, 373-388. doi: 10.1007/s10734-004-6779-5

Milliron, V., & Sandoe, K. (2008). The Net Generation Cheating Challenge. *Innovate: Journal Of Online Education*, *4*(6)

Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017-1054.

Moon, J. (2008). Critical thinking: An exploration of theory and practice. London, England: Routledge.

Moore, M. G. (1989). Three modes of interaction. *The American Journal of Distance Education, 3*(2), 1-6.

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94,* 156-163.

Murray, M., Pérez, J., Geist, D. & Hedrick, A. (2013). Student interaction with content in online and hybrid courses: Leading horses to the proverbial water. *Informing Science: The International Journal of Emerging Transdicipline, 16,* 99-114.

National Research Council (NRC). (1996). *National science education standards*. Washington, DC: National Academies Press.

National Research Council (NRC). (2005). *Measuring literacy: Performance levels for adults*. Washington, D.C: National Academies Press.

National Research Council (NRC). (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: National Academies Press

National Science Foundation (NSF), National Science Board. (2004). The message of the 2004 S & E indicators: An emerging and critical problem of the science and engineering workforce, Washington, DC: U. S. Government Press.

National Science Teacher Association (NSTA). (2011). *NSTA position statement: Quality science education and 21st century skills.* Retrieved from www.nsta.org/about/positions/21stcentury.aspx

National Science Teacher Association (NSTA). (2013). *The next generation science standards.* Retrieved from http://www.nextgenscience.org/next-generation-science-standards

Nedungadi, P., & Raman, R. (2012) A new approach to personalization: Integrating e-learning and m-learning. *Educational Technology Research and Development, 60,* 659-678. doi: 10.1007/s11423-012-9250-9

Nevid, J. S., Pastva, A., & McClelland, N. (2012). Writing-to-Learn Assignments in Introductory Psychology: Is There a Learning Benefit?. *Teaching Of Psychology*, *39*(4), 272-275.

Norris, S. P., & Phillips, L. M. (2003). How Literacy in Its Fundamental Sense Is Central to Scientific Literacy. *Science Education*, *87*(2), 224.

Norris, S. P., & Phillips, L. M. (2008). Reading as inquiry. In Duschl, R. A. & Grandy, R. E. (Eds.), *Teaching scientific inquiry: Recommendations for research and implementation.* Rotterdam, NL: Sense.

Norris, S. P., Falk, H., Federico-Agraso, M., Jimenez-Aleixandre, M. P., Phillips, L. M., & Yarden, A. (2009). Reading science texts- Epistemology, inquiry, authenticity- A rejoinder to Johnathan Osborne. *Research in Science Education, 39,* 405-410. doi: 10.1007/s11165-008-9118-5

Northrup, P. T., & Rasmussen, K. (2000). Considerations for designing web-based programs. *Computers in The Schools*, *17*(3/4), 33-46.

Northrup, P. T. (2001). A framework for designing interactivity into web-based instruction. *Educational Technology, 41*(2), 31-39.

Oksenberg, N., Stevison, L., Wall, J. D., Ahituv, N. (2013). Function and regulation of AUTS2, a gene implicated in autism and human evolution. *PLOS Genetis, 9*(1); e1003221. doi: 10.137/journal.pgen.1003221

Ondrusek, A. L. (2012). What the research reveals about graduate students' writing skills: A literature review. *Journal of Education for Library and Information Science, 55*(3), 176-188.

Organization for Economic Cooperation and Develompment (OECD). (1998). *Instrument design: A framework for assessing scientific literacy.* Report of Project Managers Meeting, Amhem, The Netherlands: Programme for International Student Assessment (PISA).

Organization for Economic Cooperation and Developmpent (OECD). (2008). *Education at a glance: OECD indicators, 2007*

Parker, K., Lenhart, A., & Moore, K. (2011). The digital revolution and higher education: College presidents, public differ on value of online learning. Washington, DC: Pew Research Center. Retrieved 4/1/14 from http://www.pewinternet.org/files/ old-media/Files/Reports/2011/PIP-Online-Learning.pdf

Partnership for 21st Century Skills (p21). (2008). *21st century skills, education & competitiveness: a resource and policy guide.* Retrieved 5/14/2014 from: http://www.p21.org/storage/documents/21st_century_skills_education_and_competitiveness_guide.pdf

Partnership for 21st Century Skills (p21). (2011). *Common Core Toolkit.* Retrieved August 23, 2013 from http://www.p21.org/tools-and-resources/p21-common-core-toolkit

Paivio, A. (1969). Concrete-image and verbal memory codes. *Journal of Experimental Psychology, 80*, 279-285. doi: 10.1037/h0027273

Paivio, A. (2013). Dual coding theory, word abstractness, and emotion: A critical review of Kousta et al. (2011). *Journal Of Experimental Psychology: General*, *142*(1), 282-287. doi: 10.1037/a0027004

Paul, R., & Elder, L. (2003). *A miniature guide for students and faculty to scientific thinking.* Dillon Beach, CA: Foundation for Critical Thinking Press.

Paul, R., & Elder, L. (2007). The thinker's guide to analytic thinking: How to take things apart and what to look for when you do. Dillon Beach, CA: Foundation for Critical Thinking Press.

Pedersen, S., & Liu, M. (2001). *The effects of modeling expert cognitive strategies during problem-based learning*. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA. Available at: http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED453264

Perkins, C. & Murphy, E. (2006). Identifying and measuring individual engagement in critical thinking in online discussions; An exploratory case study. *Educational Technology & Society, 9*(1), 298-307.

Piaget, J. (1985). *The equilibrium of cognitive structures.* Chicago, IL: University of Chicago Press.

Poniastowski, K. (2012). Getting students ready to write: An experiment in online teaching and learning. *Journalism & Mass Communication Educator, 67*(2), 120-133. doi: 10.1177/1077695812440943

Price, L., Richardson, J. T. E., & Jelfs, A. (2007). Face-to-face versus online tutoring support in distance education. *Studies in Higher Education, 32*(1). 1-20.

Quitadamo, I. J., & Kurtz, M. J. (2007). Learning to improve: Using writing to increase critical thinking performance in general education biology. *CBE Life Sciences Education, 6*(2), 140-154. doi: 10.1187/cbe.06-11-0203

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Liffiefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21-33.

Reiser, R. A. (1994). Clarke's invitation to the dance: An instructional designer's response. *Educational Technology Research & Development*, *42*(2), 45-49.

Reiser, R. A. (2001). History of instructional design and technology: Part 1: History of instructional media. *Educational Technology Research & Development, 49*(1), 53-65.

Reynolds, J., & Moskovitz, C. (2008). Calibrated peer review assignments in science courses: Are they designed to promote critical thinking and writing skills? *Journal of College Science Teaching, 38*(2), 60-66. Retrieved from http://search.proquest.com/docview/61899157?accountid=12043

Rogers, E. M. (2003). *Diffusion of innovations* (5th edition). New York, NY: Free Press.

Ross, P. M., Burgin, S., Aitchison, C., & Catterall, J. (2011). Research writing in the sciences: Liminal territory and high emotion. *Journal of Learning Design, 4*(3), p 14-27.

Saad, L., Busteed, B., & Ogisi, M. (2013). In U.S., Online education rated best for value and options: Viewed as weakest in terms of trusted grading and acceptance by employers. Retrieved from: gallup.com/poll/165425/online-education-rated-best-value-options.aspx

Samuels, S. J., & Farstrup, A. E. (2006). *What research has to say about fluency instruction.* Newark, DE: International Reading Association.

Saxton, E., Belanger, S., Becker, W. (2012). The Critical Thinking Analytic Rubric (CTAR): Investigating intra-rater and inter-rater reliability of a scoring mechanism for critical thinking performance assessments. *Assessing Writing, 17,* 251-270. Available at: http://dx.doi.org/10.1016/j.asw.2012.07.002

Scriven, M., & Paul, R. (1987). *Critical Thinking as Defined by the National Council for Excellence in Critical Thinking*, *198.* Presented at the 8th Annual International Conference on Critical Thinking and Education Reform, Summer 1987. Available at: http://www.criticalthinking.org/pages/defining-critical-thinking/766

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *elearnspace, everything elearning.* doi: 11-09-2006

Sitzmann, T., Brown, K. G., Kraiger, K., & Kanar, A. M. (2009). A multilevel analysis of the effect of prompting self-regulation in technology delivered instruction. *Personnel Psychology, 93*(4), 697-734.

Smith, J. I., & Tanner, K. (2010). The problem of revealing how students think: Concept inventories and beyond. *CBE-Life Sciences Education, 9*, 1-5. doi: 10:1187/cbe.09-12-0094

Snowman, J., & Biehler, R. (2006). *Psychology Applied to Teaching.* Boston, MA. Houghton Mifflin Company.

Sousa, D. A. (2011). *How the brain learns* (4th ed.). Thousand Oaks, CA: Corwin Press.

Stav, J. B, Nielsen, K. L., Hansen-Nygard, G., & Thorseth, T. M. (2010). Experiences obtained

with integration of student response systems for iPod touch and iPhone into e-learning environments. *Electronic Journal of e-Learning, 8*(2), 179-190. Retrieved from http://www.ejel.org/volume8/issue2/p179

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating inter-rater reliability. *Practical Assessment, Research & Evaluation*, 9(4). Retrieved from http://PAREonline.net/getvn.asp?v=9&n=4

Stouppé, I. (1998). Measuring interactivity. *Performance Improvement, 37*(9), 19-23.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition & Instruction, 8*(4), 351-363. Available at: http://www.jstor.org/stable/3233599

Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., & Cooper, S. (2006). Teaching courses online: A review of the research. *Review of Educational Research, 76,* 93-135. Retrieved from http://www.jstor.org/stable/3700584

Tardy, C. M. (2005). It's like a story: Rhetorical knowledge development in advanced academic literacy. *Journal of English for Academic Purposes, 4,* 325-338.

Tennyson, R. D., & Rasch, M. (1988). Linking cognitive learning theory to instructional prescriptions. *Instructional Science, 17*, 369-385.

Terry, D. R. (2012). Assessing critical thinking skills using articles form the popular press, *Journal of College Science Teaching, 42*(1), 66-70.

Tollefson, N. (2000). Classroom applications of cognitive theories of motivation. *Educational Psychology Review, 12*(1), 63-83.

Tsui, L. (2002). Fostering critical thinking through effective pedagogy: Evidence from four institutional case studies. *The Journal of Higher Education, 73,* 743-747.

Twigg, C. (2001). *Innovations in online learning: Moving beyond no significant difference*. Troy, NY: The Pew Learning and Technology Program.

U.S. Department of Commerce, Economics and Statistics Administration (U.S. DCESA). (2011). *STEM: Good Jobs Now and for the Future*. ESA Issue Brief #03. Retrieved August 22, 2013 from esa.doc.gov/sites/default/files/reports/documents/stemfinalyjuly14_1.pdf. T

Wang, Q., & Woo, H. L. (2007). Comparing asynchronous online discussions and face-to-face discussions in a classroom setting. *British Journal of Educational Technology, 38*(2), 272-286.

Ward, C. L. (2010). The exponential potential for different learning in online environments, *Journal of Instructional Delivery Systems, 24*(1), 8-15.

Washington State University, Center for Teaching, Learning and Technology (WSU CTLT), (2006). WSU's critical and integrative thinking rubric. Retrieved 9/24/2010 from: https://my.wsu.edu/portal/page?_pageid=177,276578&_dad=portal&_schema=PORTAL

Watkins, J., & Mazur, E. (2013). Retaining Students in Science, Technology, Engineering, and Mathematics (STEM) Majors. *Journal of College Science Teaching*, *42*(5), 36-41

Watson, G., & Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses?. *Online Journal Of Distance Learning Administration*, *13*(1). Available at: http://www.westga.edu/~distance/ojdla/spring131/watson131.html

WICHE Cooperative for Educational Technology (WCET). (2009). *Best practice strategies to promote academic integrity in online education. Version 2.0*. Available at: http://eric.ed.gov.ezproxy.lib.lehigh.edu/?id=ED536188

Webber, S., Boon, S. & Johnston, B. (2005). A comparison of UK academics' conceptions of information literacy in two disciplines: English and Marketing/" *Library and Information Research, 29*(93), 4-15.

Wenglinsky, H. (2005). *Using technology wisely: The keys to success in schools.* New York, NY: Teachers College Press.

Wiggins, G. & McTighe, J. (2005). *Understanding by Design.* Alexandria, VA.: Association for Supervision and Curriculum Development.

Woodin, T., Smith, D., & Allen, D. (2009). Transforming undergraduate biology education for all students: An action plan for the twenty-first century. *Cell Biology Education-Life Sciences Education, 8,* 271-273. doi: 10.1187/cbe.

09-09-0063

Zdravkovic, N. (2010). Spicing Up Information Literacy Tutorials: Interactive Class Activities that Worked. *Public Services Quarterly*, *6*(1), 48-64.

Zhang, D., & Zhour, L. (2003). Enhancing e-learning with interactive multimedia. *Information Resources Management Journal, 16*(4), p. 1-14.

Appendices

APPENDIX A

Critical Analysis Rubric for
Student Analysis of a Scientific Research Paper

Rater Name: ______________________ Artifact Number: _______ Average Score per Section: _______ Total Score: _______

Instructions: For each of the criteria below:

a) **Circle specific phrases** which describe the work and write comments in box below.
b) **Determine a numeric score** for each of the criteria (or indicate a half-point increment).
c) **Determine the Average Score and Total Scores** and enter on last page and in the appropriate spaces above.

1. Identifies and focuses on the main purpose and point of view of the study.

Absent	Minimal	Emerging	Developing	Competent	Effective	Mastering
0	**1**	**2**	**3**	**4**	**5**	**6**
This aspect of the paper is totally missing.	Attempts with limited success to identify either the main purpose or the point of view of the author of the paper that is being analyzed (not both). Scope is either overly narrow or overly broad.	Attempts with limited success to identify both the main purpose and the point of view of the paper being analyzed. Scope may still be overly narrow or overly broad.	Identifies and tries to summarize the main purpose of the study and the author's point of view, although some aspects may be inaccurate or confused, Several key details or nuances are missing or glossed over.	Clearly summarizes and focuses on the study's main purpose and the point of view that the author had in mind when writing the paper, though some aspects may be extraneous or inappropriately weighted. Details or nuances may still be missing or glossed over.	Accurately summarizes the study's main purpose and the author's point of view and starts to explore some of the questions, aspects or relationships identified by the author(s) with an attempt to explain complex aspects of both the paper's main focus and point of view.	Captures the complex scope of the original paper's main focus and point of view by thoroughly exploring both of these aspects and the significant underlying questions, aspects or relationships identified by the author(s).
Comments:				**Numeric Score for this Section =**		

2. Identifies and explains the most important scientific information provided and key scientific concepts that must be understood in order to analyze this paper.

Absent	Minimal	Emerging	Developing	Competent	Effective	Mastering
0	**1**	**2**	**3**	**4**	**5**	**6**
This aspect of the paper is totally missing.	Attempts to identify at least one key scientific concept that must be understood in order to analyze this paper. Scope of analysis is either overly narrow or overly broad with major inaccuracies.	Attempts with limited success to both identify and summarize the most important scientific information and key concepts that must be understood in order to analyze this paper. Scope may still be overly narrow or overly broad or some minor inaccuracies exist.	Identifies the most important scientific information needed to analyze this paper and some related key scientific concepts though minor aspects of their application to the study at hand may be inaccurate or seem confused. Details or nuances of the key concepts are missing or glossed over.	Focuses on explaining the most relevant scientific information needed to analyze this paper and any related key scientific concepts though some aspects of the explanation may be extraneous or inappropriately weighted. Some details or nuances may still be glossed over.	Fully explains the most relevant scientific information and starts to explore the significant aspects or relationships between the various key concepts needed to analyze this paper. Attempts to explain the dynamic nature and complex scope of the underlying science involved.	Thoroughly explores the most important scientific information and key concepts needed to analyze this paper with emphasis on their significant aspects and relationships. Captures the dynamic nature and complex scope of the underlying science involved in the issue at hand.
Comments:					**Numeric Score for this Section =**	

APPENDIX A

3. Identifies and considers the conclusions drawn by the author(s) and the main assumptions underlying them.

Absent	Minimal	Emerging	Developing	Competent	Effective	Mastering
0	1	2	3	4	5	6
This aspect of the paper is totally missing.	Attempts to identify the main conclusions of the study or the major assumptions being made by the author(s) of the paper being critiqued. Scope of analysis is either overly narrow, overly broad or contains major inaccuracies.	Attempts with limited success to identify both the main conclusions of the study and the assumptions of the scientific arguments made by the author(s) of the paper being critiqued. Scope may still be overly narrow or overly broad.	Identifies and starts summarizing main conclusions and any underlying assumptions of the author(s) and starts to summarize them although minor aspects may be inaccurate or confusing. Key details and/or nuances are missing or glossed over.	Summarizes and starts to explain the conclusions of the study and the underlying assumptions made by the authors of the paper being analyzed, Some aspects of the summary and/or explanation may be extraneous or inappropriately weighted. Some details or nuances may still be glossed over.	Cursory exploration of the most important conclusions and assumptions made by the author(s) including how they are related to the main science issues related to the study being analyzed. Starts to express the dynamic nature and complex scope of the study at hand..	Thoroughly explores the most important scientific information and key concepts needed to analyze this paper with emphasis on their significant aspects and relationships. Captures the dynamic nature and complex scope of the underlying science involved in the issue at hand.

Comments:

Numeric Score for this Section =

APPENDIX A

4. Identifies the main implications of this paper if the findings of this study are later accepted or rejected by others.

Absent	Minimal	Emerging	Developing	Competent	Effective	Mastering
0	**1**	**2**	**3**	**4**	**5**	**6**
This aspect of the paper is totally missing.	Attempts to identify at least one main implication of either accepting or rejecting the findings of this study. No mention of whether the student accepts or rejects the findings of this study. Scope is overly narrow or too broad.	Student attempts with limited success to identify and summarize more than one implication of accepting and/or rejecting the findings of this study without taking a clear personal stand on the results of the study. Scope may still be overly narrow or overly broad.	Student clearly accepts or rejects the main findings of the study and includes a cursory discussion of the main implications of this decision with some inaccurate or confusing statements and no mention of how they made their decision. Key details are missing and some nuances glossed over.	Starts to explain why they chose to reject or accept the findings of the study at hand discussing both options. Some aspects of the summary and/or the explanation may be extraneous or inappropriately weighted. Some details or nuances may still be missing or vague.	Explains the reasoning behind choosing to accept or reject the findings of the study being analyzed and the reasons they made their decision. Starts to explore the implications of both accepting the main findings of the study and of rejecting them without discussing ways to resolve dilemmas nor making suggestions for future studies in the field..	Clearly and thoroughly explains the implications of both accepting and of rejecting the findings of this particular scientific paper. Student fully justifies their stance and takes responsibility for any resulting dilemmas created by their decision to reject or accept the major findings of the study at hand by suggesting appropriate avenues for future research.
Comments: **Numeric Score for this Section =**						

APPENDIX A

5. Analyzes the credibility of the sources used by the author(s) in researching the topic at hand.

Absent	Minimal	Emerging	Developing	Competent	Effective	Mastering
0	**1**	**2**	**3**	**4**	**5**	**6**
This aspect of the paper is totally missing.	Contains only a passing mention that the author(s) cited other scientific sources in the paper being analyzed.	Mentions the importance of a thorough review of the existing scientific literature or lists the total number of citations used when planning this study as an indicator of thoroughness	Discusses the complexity of scientific research and the impact of a high quality scientific literature review process on the overall perception of the quality of the study being analyzed.	Mentions crude ways of measuring the quality of the author(s)' initial literature review when planning this study, but fails to include key criteria used in judging reliability of the resources themselves.	Lists criteria used to judge the reliability of individual resources cited in a literature review and makes an initial judgment of the overall quality of the literature review process used by the author(s) of the analyzed paper.	Critiques the reliability of the resources cited in the analyzed paper using the pre-established criteria and uses these critiques to make a clear statement about the overall reliability of the study at hand.
Comments:				**Numeric Score for this Section =**		

Section	1-Purpose & Point of View	2-Key Scientific Concepts	3-Conclusions & Assumptions	4-Implications	5-Cited Sources	Total Score
Score						
Comments:				**Average Section Score =**		

Total Score and Average Score entered on this page and also in the appropriate field on the first page of this rubric.

Thank you.

APPENDIX B

Assignment Instructions for First Paper to Analyze - Group A only

Visit **http://tinyurl.com/paperAgenetics** to access a copy of the first scientific study you will be analyzing. After reading this paper, write an analysis between 700-800 words in length. Please use an objective voice and, for full credit, include these five critical analysis elements:

- Identify the main purpose of the study being analyzed and point of view of the authors.
- Identify and discuss the most important scientific information provided and any key scientific concepts that must be understood in order to analyze this original study.
- Identify and discuss conclusions drawn by the researchers/authors and list the main assumptions underlying them.
- Identify the main implications of this scientific article if the findings of this study are accepted or rejected. Then share your personal opinion of the research.
- Analyze the credibility of the sources used by the authors in researching the topic at hand. This will require an overview of the references the researchers used both for the literature review and in planning their methodology.

Use at least two reliable sources to verify the information in this paper, preferably other primary scientific research articles that have been published in peer-reviewed journals. Review articles summarize multiple primary research studies, but may not provide all of the details needed to analyze the basic premises of specific experimental studies. The original paper being analyzed should be cited in the reference section of the analysis paper along with any additional references used. Use MLA format for all citations (see www.mla.org).

This assignment uses *Turnitin*, a software program designed to help you practice using and properly citing the work of others without inadvertently plagiarizing it. Each time you submit a draft of your paper, *Turnitin* will analyze it and provide you with an originality score. The more you explain the paper in your own words, the better your originality score. You can replace multiple drafts and update your originality score up until the assignment deadline.

Finally, the paper should be double-spaced using a 12-pt font with one-inch margins all around. For assistance with citations and related research topics, visit LU libraries at:

http://library.lehigh.edu/teaching_support/information_literacy_teaching_research_skills/

The final draft of this first paper must be submitted here by midnight on Sunday, October 19, 2014. No late submissions will be accepted. Students who are participating in the study will then receive instructions on accessing additional information on analyzing scientific papers.

APPENDIX C

Critical Analysis Pre-test

1) What is a review article and how is it related to a primary research article?

2) Name one way to emphasize information in scientific writing.

3) Where is study data usually found and why?

4) Which sources need to be cited in the reference section of your analysis paper?

- ○ All of the sources you used in your research
- ○ Only the sources you have directly quoted
- ○ Only books and magazines, but not websites
- ○ None of the sources you used in your research
- ○ Only the article being analyzed
- ○ All of your sources except for the article being analyzed

5) The first section of a primary scientific article is the

- ○ Introduction
- ○ Methods
- ○ Title
- ○ Abstract

APPENDIX C

6) Sources used by the authors to plan their study can be found in the _________ section.

- ○ References
- ○ Introduction
- ○ Results
- ○ Discussion
- ○ Methods

7) The strength of a scientific study lies in:

- ○ Using a proven theory
- ○ How well extraneous variables were controlled
- ○ Including tables of raw data
- ○ Proven hypotheses

8) **Your Background** - please circle the correct answer or fill in the blanks:

Gender: Male Female

Age: ___________________

Intended major/career:

____

APPENDIX C

Total number of college credits completed to date: __________

Current GPA _________

Number college science or engineering courses completed to date: __________

Total number science credits: _______ Total number engineering credits: _______

Number of college writing courses completed to date: ________

Did you get AP credit in English? Yes No

If so, how many credits worth? ______________

Number of college history courses completed to date: ________

Please rate yourself on the following skills using a scale from 5 (excellent) to 1 (poor):

	Very Good	good	average	below ave.	poor	Not/Applicable
Scientific reader	5	4	3	2	1	N/A
Scientific writer	5	4	3	2	1	N/A
Research skills	5	4	3	2	1	N/A
Science Critic	5	4	3	2	1	N/A
Information Literacy	5	4	3	2	1	N/A
Scientific Literacy	5	4	3	2	1	N/A

Thank you for your time

APPENDIX D

LEHIGH UNIVERSITY

Consent Form

Online Instruction in Critical Analysis Skills: Focused reading and Science Writing

Your instructor has agreed to participate in a research study of how different online instructional approaches can help students develop better critical analysis (CA) skills. The primary investigator would like to measure the level of CA skills demonstrated within scientific analysis papers you will be writing anyway for an assignment required by your course instructor, but will need your express permission to so. If you decide to participate in this study, several trained educators who are not at all affiliated with Lehigh University will blindly rate the CA skills found within anonymous versions of the papers you submit for a course grade. All personal information about you will be removed from these papers and be replaced with a study code. The code database will be available only to the primary investigator. Your instructor will not be involved in the rating, nor even know which students chose to participate in the study. The study results will not be able to impact your grade because: (a) only group data will be reported and (b) results will be reported only after the grade submission deadline for the semester. No one, including your instructor and the raters will be able to connect any CA results to a particular student. Please read this form carefully and ask the investigator questions you may have before agreeing to be part of this study.

This study is being conducted by: Cindy Chesworth Adams, a doctoral student in the Learning Sciences and Technology (LST) doctoral program at Lehigh University.

APPENDIX D

Cindy is working under the direction of Dr. Lynn Columba, Associate Professor, Teaching, Learning and Technology (TLT) program, College of Education, Lehigh University, Bethlehem, PA.

Individual CT results will be kept strictly confidential and course instructors will see aggregated group data only. No data will be given to instructors in any form until well after course grades are submitted for the semester under study (see section on confidentiality below) so there is no way that your decision to participate (or not) will affect your course grade in any way. Should anyone become concerned about any aspect of this study, they will be referred to the Lehigh University Counseling Center for assessment and follow-up if needed. These services are free of charge to all current Lehigh University students.

The benefits to participation in this study are:

1) The opportunity to discover your individual CA skill score. If you are interested in knowing this, you will need to send an e-mail request to the primary researcher, Cindy Chesworth Adams (see contact information on the back of this form) by December 31, 2014.
2) The opportunity to help science faculty at Lehigh determine the best ways to help students enhance CA, academic writing and critical thinking skills while at Lehigh University. These skills are important to all students in their future STEM (Science, technology, engineering or mathematics) careers and also in their roles as scientifically literate citizens and voters.

You will not receive payment for your participation. However, all students who sign a consent form to allow use of their paper in this study will be entered into two drawings for one of five gift certificates to the LU bookstore worth $25 apiece.

The records of this study will be kept confidential and any information collected through this research project that personally identifies you will not be voluntarily released or disclosed without your separate consent, except as specifically required by law. In any sort of report we might publish, we will not include any information that will make it possible to identify an individual subject. Research records will be stored securely and only the primary researchers will have access to the records.

Supplemental online instruction, including one of three different interventions, will be presented through CourseSite, Lehigh University's LMS, and will be monitored and graded objectively by the researcher. Your posted paper and comments will be reviewed for CT content as part of this study.

Participation in this study is totally voluntary: Your decision whether or not to participate will not affect your current or future grade, relations with Lehigh University of any of your course instructors. If you decide to participate, you are free to choose not to answer any specific questions or to withdraw at any time from the study without affecting any of the relationships listed here.

The primary researcher conducting this study is: Cindy Chesworth Adams, MS, LST Doctoral Candidate at Lehigh University. Any questions you may have about the study can be directed first to Ms. Adams at 484-695-2879 (e-mail ccadams@lehigh.edu) or her faculty advisor, Dr. Lynn Columba, who can be reached via e-mail at hlc0@lehigh.edu (office phone number is 610-758-3237.

If you have any questions or concerns regarding this study and would like to talk to someone other than the researchers, **you are encouraged** to contact Susan Disidore at (610) 758-3021 (email: inors@lehigh.edu) of Lehigh University's Office of Research and Sponsored Programs. All contact with this office or any reports or correspondence will be kept strictly confidential.

I have read the above information and understand the study described above and I have been given a copy of the study description. I am 18 years of age or older and I agree to participate in this study.

Print Name____

Signature:__________________________________ Date: __________

Signature of Investigator:__*Cynthia C. Adams*___ Date: _08/25/2014__

Your electronic signature indicates your permission to participate in this study.

Please print a copy of the signed form for your records prior to submitting it electronically through the *Qualtrics* survey site for this study.

APPENDIX E

Critical Analysis Post-test

1. What is a review article and how is it related to a primary research article?

2. Where is study data usually found and why?

3. Which sources need to be cited in the reference section of your analysis paper?

   - All of the sources you used in your research
   - Only the sources you have directly quoted
   - Only books and magazines, but not websites
   - None of the sources you used in your research
   - Only the article being analyzed
   - All of your sources except for the article being analyzed

4) The results of a scientific study would not be found in the:

   - Abstract
   - Introduction
   - Tables and Figures
   - Discussion

5) The sources used by the authors to plan their study can be found in the _________ section.

   - References

- Introduction
- Results
- Discussion
- Methods

6) The strength of a scientific study lies in:

- Using a proven theory
- How well extraneous variables were controlled
- Including tables of raw data
- Proven hypotheses

7) Please rate yourself on the following skills using a scale from 5 (excellent) to 1 (poor):

	Very Good	good	average	below ave.	poor	Not/Applicable
Scientific reader	5	4	3	2	1	N/A
Scientific writer	5	4	3	2	1	N/A
Research skills	5	4	3	2	1	N/A
Science Critic	5	4	3	2	1	N/A
Information Literacy	5	4	3	2	1	N/A
Scientific Literacy	5	4	3	2	1	N/A

Please rate the following statements about this study using a scale from 5 (excellent) to 1 (poor):

	Very Good	good	average	below ave.	poor	Not/Applicable
Study purpose was clear	5	4	3	2	1	N/A
Online lesson was helpful	5	4	3	2	1	N/A

Incentives were attractive	5	4	3	2	1	N/A
Overall study was helpful	5	4	3	2	1	N/A

Would you participate in this type of study again? Why or why not?

How could this study have been improved?

What other kind of resources would you like to see for scientific communication?

Any other comments?

Thank you for your time

Cynthia Chesworth Adams

Biography

Originally trained as a scientist, Cindy holds a B.S. degree in biology from the Pennsylvania State University where she completed a senior thesis in Food Science. She went on to earn a M.S. degree in nutritional biochemistry from Rutgers, the State University of New Jersey where she was a teaching assistant and a stress management peer educator. Her first professional job was as a community health educator in Morris County, NJ. Certified as a Health Education Specialist (CHES) in 1990, she has over 20 years of experience as a health educator in corporate, community and hospital settings. She was the Health Promotion Coordinator at Lafayette College from 1994-2004.

From 2007-2010 she oversaw the development of the biomanufacturing curricula and the construction of a cleanroom training facility at Northampton Community College. (NCC). She has written or co-written grants totaling over $900,000 during her career

She completed a six-month doctoral internship at Lehigh Valley Hospital in 2011 that led to full-time employment as an Instructional Designer for an electronic health record software company. Cindy is currently an adjunct professor of anatomy and physiology (A&P) at both NCC and East Stroudsburg University. She started teaching health online at NCC in January of 2009 and is currently teaching a hybrid A&P course that meets face-to-face once a week for anatomy laboratory sessions at NCC.

EDUCATION

Lehigh University Bethlehem, PA	Ph.D. in Learning Sciences and Technology	2015
Rutgers University New Brunswick, NJ	M.S. in Nutritional Biochemistry	1983
Penn State University University Park, PA	B.S. in Biology	1981

PUBLICATIONS

Adams, C. & Columba, L. (2014). Classroom response systems: Effects on the critical analysis skills of students in introductory science courses. *School Science and Mathematics, 114*(8), 367-379.

Adams, C., & Bortz, C. (2010). The PETRI Project Pipeline for Education, Training Resources, and Innovation. *Community College Journal of Research and Practice, 34,* 11, 954-956.